138758

STIC-Biotech/ChemLib

From: Chan, Christina
Sent: Friday, November 26, 2004 3:51 PM
To: Schnizer, Richard; STIC-Biotech/ChemLib
Subject: RE: 10/798,532 RUSH

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----
From: Schnizer, Richard
Sent: Friday, November 26, 2004 7:14 AM
To: Chan, Christina
Subject: 10/798,532 RUSH

Please authorize the following RUSH search request:

Please search the commercial and interference databases for SEQ ID NOS: 1-4 from 10/798,532. Please also search for nucleic acids that could encode the polypeptides of SEQ ID NOS: 2 and 4.

Thank you-

Richard Schnizer, Ph.D.
Patent Examiner
Art Unit 1635
Remsen 2D30
571-272-0762
Mail Box 2C18

****
STAFF USE ONLY

Searcher: Noble
Searcher Phone: 2-
Date Searcher Picked up:
Date Completed: Nov 30 2004
Searcher Prep/Rev. Time: 5
Online Time: 3

****
Type of Search
NA Sequence: # 4
AA Sequence :# 2
Structure: #
Bibliographic:
Litigation:
Patent Family:
Other:

****
Vendors and cost where applicable
STN:
DIALOG:
QUESTEL/ORBIT:
LEXIS/NEXIS:
SEQUENCE SYSTEM: COMPUGEN
WWW/Internet:
Other(Specify):

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:30 ; Search time 5310 Seconds
(without alignments)
8646.716 Million cell updates/sec

Title: US-10-798-532-1
Perfect score: 1260
Sequence: 1 atggcgggatccggctgcgc.....tgtatcaaaccttgctttaa 1260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1115	88.5	1763	3	AK013347	AK013347 Mus muscu
2	886.6	70.4	2846	3	AK030152	AK030152 Mus muscu
3	861.8	68.4	1079	4	BM468107	BM468107 AGENCOURT
4	723.8	57.4	917	4	BM451184	BM451184 AGENCOURT
5	624.4	49.6	1015	6	BY712344	BY712344 BY712344
6	619.2	49.1	657	4	BG547675	BG547675 602575520
7	617.8	49.0	899	2	BE895119	BE895119 601436041
8	616.4	48.9	772	4	BI562420	BI562420 603256316
9	610.4	48.4	932	5	BU541306	BU541306 AGENCOURT
10	595	47.2	875	4	BI259568	BI259568 602968330
11	593.2	47.1	834	7	CF593363	CF593363 AGENCOURT
12	585.2	46.4	682	4	BI333581	BI333581 602997069
13	583.8	46.3	906	4	BG386905	BG386905 602454704
14	579.6	46.0	888	5	BQ225071	BQ225071 AGENCOURT
15	560	44.4	602	7	CO737403	CO737403 SlLT02c17
16	558	44.3	876	5	BU184346	BU184346 AGENCOURT
17	557.6	44.3	779	5	BU447831	BU447831 603767571
18	550.8	43.7	844	4	BG474616	BG474616 602517423
19	519.8	41.3	824	2	BE568528	BE568528 601342425
20	516.8	41.0	823	7	CO771206	CO771206 testis_ES
21	516	41.0	781	7	CN232453	CN232453 WLB074D11
22	499.4	39.6	670	7	CN359673	CN359673 170004245
23	498.6	39.6	717	5	BU385761	BU385761 603582650
c 24	496.8	39.4	523	2	BF929410	BF929410 IL2-NT020
25	495.2	39.3	896	7	CO773091	CO773091 testis_ES
26	494.6	39.3	928	5	BQ424438	BQ424438 AGENCOURT
27	494.2	39.2	1003	5	BQ922020	BQ922020 AGENCOURT
28	482.4	38.3	716	7	CF745774	CF745774 UI-M-GV0-
29	471.8	37.4	755	5	BU371858	BU371858 603568233
30	469.8	37.3	794	5	BU205357	BU205357 603104569
31	468.8	37.2	785	6	CB990192	CB990192 AGENCOURT
32	468.2	37.2	824	4	BI253163	BI253163 602973284
33	465	36.9	699	5	BU707284	BU707284 UI-M-FR0-
34	463.6	36.8	870	5	BQ216813	BQ216813 AGENCOURT
35	463.4	36.8	484	5	BX281211	BX281211 BX281211
36	460	36.5	696	5	BU439278	BU439278 604147528
c 37	456.6	36.2	483	2	BF929411	BF929411 IL2-NT020
38	454.2	36.0	631	5	BU112717	BU112717 603131421
39	451.8	35.9	743	4	BI562440	BI562440 603256324
40	450.6	35.8	933	2	BE886040	BE886040 601507192
c 41	450.4	35.7	986	5	BU751496	BU751496 CH4#002_E
42	433.2	34.4	820	7	CK597557	CK597557 AGENCOURT
43	427.4	33.9	672	5	BX261106	BX261106 BX261106
44	420	33.3	604	5	BX261107	BX261107 BX261107
45	413.2	32.8	570	5	BU293329	BU293329 604166812

ALIGNMENTS

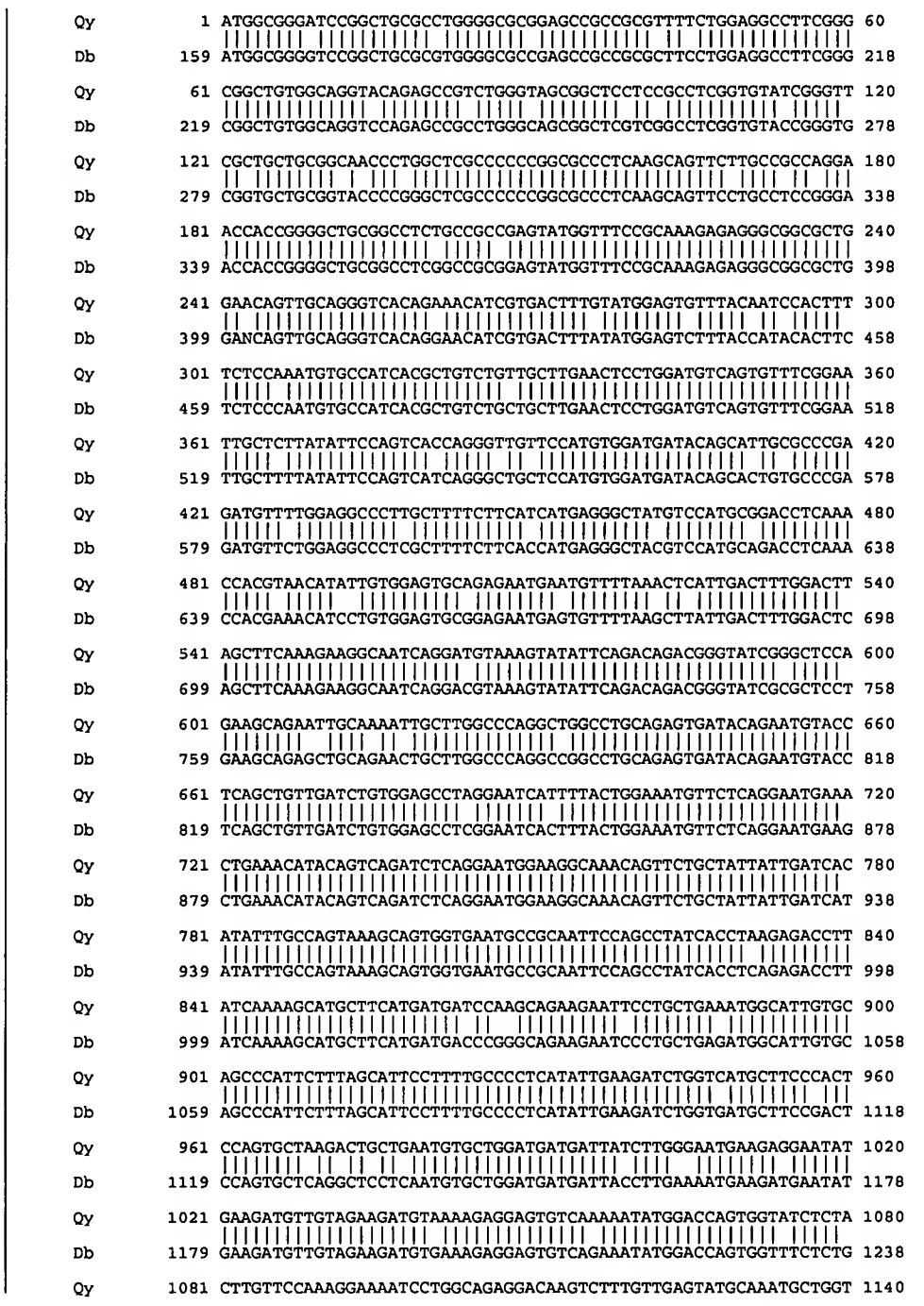
RESULT 1
AK013347
LOCUS AK013347 1763 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810454O04 product:kinase interacting with leukemia-associated gene (stathmin), full insert sequence.
ACCESSION AK013347
VERSION AK013347.1 GI:12850651
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1763)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES Location/Qualifiers
source 1. .1763
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2810454O04"
/db_xref="taxon:10090"
/clone="2810454O04"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10, 11 days embryo"
CDS 159. .1418
/note="unnamed protein product; kinase interacting with leukemia-associated gene (stathmin) (MGD|MGI:1341908) putative"
/codon_start=1
/protein_id="BAB28802.1"
/db_xref="GI:12850652"
/translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALXQLQGHRNIVTLYGVFTIHFSP NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE CTSAVDLWSLGITLLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR MFDGKFVVATFYPLSAYKRGYLYQTLL"

ORIGIN

Query Match 88.5%; Score 1115; DB 3; Length 1763;
Best Local Similarity 92.8%; Pred. No. 0;
Matches 1169; Conservative 0; Mismatches 91; Indels 0; Gaps 0;



```
                 ||||||||||||||||||||||||||||||||||||||||| |||||||| || || ||||||
Db       1239 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1298

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
                 ||||||||||||||| ||||| || |||||||| |||||||||| ||||||||||||||
Db       1299 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1358

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
                 ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1359 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1418


RESULT 2
AK030152
LOCUS       AK030152               2846 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4933401D07 product:kinase interacting with
            leukemia-associated gene (stathmin), full insert sequence.
ACCESSION   AK030152
VERSION     AK030152.1  GI:26326136
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 2846)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.jp/
            URL:http://fantom.gsc.riken.jp/.
FEATURES             Location/Qualifiers
     source          1. .2846
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4933401D07"
                     /db_xref="taxon:10090"
                     /clone="4933401D07"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             238. .1230
                     /note="unnamed protein product; kinase interacting with
                     leukemia-associated gene (stathmin) (MGD|MGI:1341908)
                     putative"
                     /codon_start=1
                     /protein_id="BAC26812.1"
                     /db_xref="GI:26326137"
                     /translation="MTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMW
                     MIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY
                     IQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQE
                     WKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALCSPFFSIP
                     FAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSLLVPKE
                     NPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL"
     polyA_signal    2823. .2828
                     /note="putative"
     polyA_site      2846
                     /note="putative"
ORIGIN

  Query Match             70.4%;  Score 886.6;  DB 3;  Length 2846;
  Best Local Similarity   93.1%;  Pred. No. 1.8e-248;
  Matches  928;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy        264 CATCGTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTG 323
                | || ||||||| |||||||| ||||| || ||||| ||||| ||||||||||||||||||
Db        234 CTTCATGACTTTATATGGAGTCTTTACCATACACTTCTCTCCCAATGTGCCATCACGCTG 293

Qy        324 TCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCA 383
              |||| |||||||||||||||||||||||||||||||||||||| |||||||||||||| ||
Db        294 TCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTTTTATATTCCAGTCATCA 353

Qy        384 GGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGAGATGTTTTGGAGGCCCTTGCTTT 443
              ||| || |||||||||||||||||||| || |||||||||||| |||||||||| |||||
Db        354 GGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGAGATGTTCTGGAGGCCCTCGCTTT 413

Qy        444 TCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGC 503
              |||||| |||||||||| |||||||| |||||||||||||| |||||  ||||||||||
Db        414 TCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAACCACGAAACATCCTGTGGAGTGC 473

Qy        504 AGAGAATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGA 563
              |||||||| |||||||| || |||||||||||||| ||||||||||||||||||||||||
Db        474 GGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTCAGCTTCAAAGAAGGCAATCAGGA 533
```

```
Qy        564 TGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTT 623
              ||||||||||||||||||||||||||||||| ||||| |||||||| |||| || |||||
Db        534 CGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCTGAAGCAGAGCTGCAGAACTGCTT 593

Qy        624 GGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCT 683
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCT 653

Qy        684 AGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCA 743
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        654 CGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAGCTGAAACATACAGTCAGATCTCA 713

Qy        744 GGAATGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGT 803
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        714 GGAATGGAAGGCAAACAGTTCTGCTATTATTGATCATATATTTGCCAGTAAAGCAGTGGT 773

Qy        804 GAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 863
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        774 GAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTTATCAAAAGCATGCTTCATGATGA 833

Qy        864 TCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTT 923
              ||  |||||||||| |||||||| |||||||||||||||||||||||||||||||||||||
Db        834 CCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTT 893

Qy        924 TGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGT 983
              |||||||||||||||||||||||| |||||||| |||||||||| || || || |||||
Db        894 TGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACTCCAGTGCTCAGGCTCCTCAATGT 953

Qy        984 GCTGGATGATGATTATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAA 1043
              ||||||||||||||| |||| |||||||| |||||||||||||||||||||||||||| ||
Db        954 GCTGGATGATGATTACCTTGAAAATGAAGATGAATATGAAGATGTTGTAGAAGATGTGAA 1013

Qy       1044 AGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGG 1103
              |||||||||||  |||||||||||||||||| |||| |||||||||||||||||||||||
Db       1014 AGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTGCTTGTTCCAAAGGAAAATCCTGG 1073

Qy       1104 CAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATT 1163
              |||||||||||||| |||||||| || || |||||||||||||||||||| ||||| ||
Db       1074 CAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGTGATTCCAAAGCTGCTCAGAAGTT 1133

Qy       1164 ACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGC 1223
              |||||||| |||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       1134 GCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGC 1193

Qy       1224 CTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||| ||||||||||||||||||
Db       1194 CTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1230
```

```
RESULT 3
BM468107
LOCUS       BM468107                1079 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6432293 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5535653
            5', mRNA sequence.
ACCESSION   BM468107
VERSION     BM468107.1  GI:18517149
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12224  row: b  column: 06
            High quality sequence stop: 754.
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5535653"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN
```

Query Match 68.4%; Score 861.8; DB 4; Length 1079;
Best Local Similarity 98.8%; Pred. No. 2.6e-241;
Matches 888; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

```
Qy        267 CGTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCT 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCT 60

Qy        327 GTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGG 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGG 120

Qy        387 TTGTTCCATGTGGATGATACAGCATTGCGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCT 446
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        121 TTGTTCCATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCT 180

Qy        447 TCATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGA 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGA 240

Qy        507 GAATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGT 566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGT 300

Qy        567 AAAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGC 626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGC 360

Qy        627 CCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGG 686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGG 420

Qy        687 AATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGA 746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGA 480

Qy        747 ATGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAA 806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAA 540

Qy        807 TGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCC 866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCC 600

Qy        867 AAGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGC 926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGC 660

Qy        927 CCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCT 986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCT 720

Qy        987 GGATGATGATTATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGA 1046
              |||||||||||||||| |||||||||||||||||||||||||||||||     |||||
Db        721 GGATGATGATTATCTTGAGAATGAAGAGGAATATGAAGATGTTGTAGAAGAATGTAAAGA 780
```

```
Qy       1047 GGAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAG 1106
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        781 GGAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGNTCCAAAGGAAAATCCTGGCAG 840

Qy       1107 AGGACAAGTC-TTTGTTGAGTATGCAAATGCTG-GTGATTCCAAAGCTGCGCAGAAATT 1163
             |||||||||| |||| |||||||||||||||| ||||||| ||||||||||||||||||
Db        841 AGGACAAGTCTTTTGNTGAGTATGCAAATGCTGNGTGATTCAAAAGCTGCGCAGAAATT 899


RESULT 4
BM451184
LOCUS       BM451184                 917 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6392719 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5495159
            5', mRNA sequence.
ACCESSION   BM451184
VERSION     BM451184.1  GI:18500224
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12121  row: j  column: 24
            High quality sequence stop: 718.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5495159"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             57.4%;  Score 723.8;  DB 4;  Length 917;
  Best Local Similarity   99.7%;  Pred. No. 8.7e-201;
  Matches  725;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        534 TGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCG 593
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 TGGGCTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCG 80

Qy        594 GGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGA 653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGA 140

Qy        654 ATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 200

Qy        714 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 260

Qy        774 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 320

Qy        834 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 380

Qy        894 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 440

Qy        954 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGA 1013
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        441 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGA 500

Qy       1014 GGAATATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGT 1073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 GGAATATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGT 560

Qy       1074 ATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAA 1133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 ATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAA 620

Qy       1134 TGCTGGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTT 1193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 TGCTGGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTT 680

Qy       1194 TGTTGTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTT 1253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 TGTTGTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTT 740

Qy       1254 GCTTTAA 1260
              |||||||
Db        741 GCTTTAA 747


RESULT 5
BY712344
LOCUS       BY712344                 1015 bp    mRNA    linear   EST 17-DEC-2002
DEFINITION  BY712344 RIKEN full-length enriched, 10, 11 days embryo whole body
            Mus musculus cDNA clone 2810454O04 5', mRNA sequence.
ACCESSION   BY712344
VERSION     BY712344.1  GI:27123621
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1015)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
```

```
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
  MEDLINE   22354683
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
            Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
            Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
            Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
            Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
            Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .1015
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="2810454O04"
                     /tissue_type="whole body"
                     /dev_stage="10, 11 days embryo"
                     /clone_lib="RIKEN full-length enriched, 10, 11 days embryo
                     whole body"
ORIGIN

  Query Match             49.6%;  Score 624.4;  DB 6;  Length 1015;
  Best Local Similarity   90.2%;  Pred. No. 1.4e-171;
  Matches  679;  Conservative    0;  Mismatches   72;  Indels    2;  Gaps    1;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              |||||||| |||||||||||| |||||||| |||||||||| || ||||||||||||||
Db        159 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 218

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              |||||||||||||| |||||||| ||||| |||||||| || |||||||||| |||||
Db        219 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 278

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              || |||||||| | ||| ||||||||||||||||||||||||||||| |||| || |||
Db        279 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 338

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||
```

```
Db        339 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 398

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              || |||||||||||||||| |||||||||||| |||||||| ||||| || |||||
Db        399 GANCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 458

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||| |||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        459 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 518

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||| |||||||||||||| ||||| || |||||||||||||||||||| || ||||||
Db        519 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 578

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||| |||||||||| |||||||||| |||||||||| |||||||| |||||||||
Db        579 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 638

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||| |||||  |||||||||| |||||||| |||||||| || ||||||||||||||
Db        639 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 698

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||| |||||||||||||||||||||||||||||| |||||
Db        699 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 758

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              |||||||| |||| || |||||||||||| |||||||||||||||||||||||||||||
Db        759 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 818

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||||  | | || ||||||| ||||||||| ||||||||||||||||
Db        819 TCAGCTGTTGATCTGTCGGGGCTCGGAATCACTTTACTGGAGATGTTCTCAGGAATGAAG 878

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAG 753
              |||||  |      |||||||||||||  ||| |
Db        879 CTGAA--AATACATCAGATCTCAGGATGGGAGG 909


RESULT 6
BG547675
LOCUS       BG547675                 657 bp    mRNA    linear   EST 04-APR-2001
DEFINITION  602575520F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703645 5',
            mRNA sequence.
ACCESSION   BG547675
VERSION     BG547675.1  GI:13546340
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1542  row: g  column: 06
            High quality sequence stop: 657.
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4703645"
                     /lab_host="DH10B (T1 phage-resistant)"
```

/clone_lib="NIH_MGC_77"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 49.1%; Score 619.2; DB 4; Length 657;
Best Local Similarity 99.1%; Pred. No. 4.2e-170;
Matches 654; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

```
Qy        577 CAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGC 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGC 60

Qy        637 CTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTA 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTA 120

Qy        697 CTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCA 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCA 180

Qy        757 AACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATT 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATT 240

Qy        817 CCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGA 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGA 300

Qy        877 ATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATT 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATT 360

Qy        937 GAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGAT 996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGAT 420

Qy        997 TATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAA 1056
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TATCTTGAGAATGAAGAGGAATATGAAGATGTCGTAGAAGATGTAAAAGAGGAGTGTCAA 480

Qy       1057 AAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTC 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTC 540

Qy       1117 TTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGG 1176
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTTG-TGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGG 599

Qy       1177 ATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGA 1236
              |||||||||||||||| | ||||||||||||||||||||||||||||| ||||||||||||
Db        600 ATGTTTGATGGGAAGTTAG-TGTGGCTACATTCTACCCGCTGAGTG-CTACAAGAGGGGA 657
```

RESULT 7
BE895119
LOCUS BE895119 899 bp mRNA linear EST 20-OCT-2000
DEFINITION 601436041F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920939 5', mRNA sequence.
ACCESSION BE895119
VERSION BE895119.1 GI:10358193
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 899)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9753 row: b column: 12
High quality sequence stop: 626.
FEATURES Location/Qualifiers
source 1. .899
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3920939"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 49.0%; Score 617.8; DB 2; Length 899;
Best Local Similarity 99.7%; Pred. No. 1.2e-169;
Matches 619; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        337 CTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATG 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATG 60

Qy        397 TGGATGATACAGCATTGCGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAG 456
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 TGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAG 120

Qy        457 GGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGT 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGT 180

Qy        517 TTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATT 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATT 240

Qy        577 CAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGC 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGC 300

Qy        637 CTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTA 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTA 360

Qy        697 CTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCA 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCA 420

Qy        757 AACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATT 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATT 480

Qy        817 CCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGA 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGA 540

Qy        877 ATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATT 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        541 ATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATT 600

Qy        937 GAAGATCTGGTCATGCTTCCC 957
              ||||||||||||||||||  ||
Db        601 GAAGATCTGGTCATGCTTTCC 621


RESULT 8
BI562420
LOCUS       BI562420                772 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603256316F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
            mRNA sequence.
ACCESSION   BI562420
VERSION     BI562420.1  GI:15449747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 772)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11755  row: p  column: 03
            High quality sequence stop: 727.
FEATURES             Location/Qualifiers
     source          1. .772
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5298674"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTVN-3', size-selected for  average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation).  Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN

  Query Match             48.9%;  Score 616.4;  DB 4;  Length 772;
  Best Local Similarity   99.4%;  Pred. No. 2.9e-169;
  Matches  628;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        274 TTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTGTTGCTT 333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 TTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTGTTGCTT 167

Qy        334 GAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCC 393
              |||||||||||||||||||||||| |||||||||||||||||| |||||||||  |||||
Db        168 GAACTCCTGGATGTCAGTGTTTCGGAANTGCTCTTATATTCCAGTCACCANGGTTGTTCC 227

Qy        394 ATGTGGATGATACAGCATTGCGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCAT 453
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        228 ATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCAT 287

Qy        454 GAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAA 513
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAA 347

Qy        514 TGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 573
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 TGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 407

Qy        574 ATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCT 633
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 ATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCT 467

Qy        634 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 693
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 527

Qy        694 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 587

Qy        754 GCAAACAGTTCTGCTATTATTGATCACA-TATTTGCCAGTAAAGCAGTGGTGAATGCCGC 812
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        588 GCAAACAGTTCTGCTATTATTGATCACATTATTTGCCAGTAAAGCAGTGGTGAATGCCGC 647

Qy        813 AATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAG 872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 AATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAG 707

Qy        873 AAGAATTCCTGCTGAAATGGCATTGTGCAGCC 904
              ||||||||||||||||||||||||||||||||
Db        708 AAGAATTCCTGCTGAAATGGCATTGTGCAGCC 739


RESULT 9
BU541306
LOCUS       BU541306                932 bp    mRNA    linear   EST 13-SEP-2002
DEFINITION  AGENCOURT_10327419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6572970
            5', mRNA sequence.
ACCESSION   BU541306
VERSION     BU541306.1  GI:22851747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 932)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2765  row: o  column: 18
            High quality sequence stop: 631.
FEATURES             Location/Qualifiers
     source          1. .932
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6572970"
                     /tissue_type="carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_40"
                     /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed  by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California,  Berkeley) using ZAP-cDNA synthesis kit
```

(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
ORIGIN

Query Match 48.4%; Score 610.4; DB 5; Length 932;
Best Local Similarity 93.9%; Pred. No. 1.8e-167;
Matches 690; Conservative 0; Mismatches 37; Indels 8; Gaps 5;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 208

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 268

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 328

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 388

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 448

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 508

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||| ||||||
Db        509 TTGCTCTTATATTCCAGTCACCAGCGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 568

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCT-CAA 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        569 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCCAA 628

Qy        480 ACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGAC- 538
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        629 ACCACGTAACATATTGTGGAGTGCAAAGAATGAATGTTTTAAACTCATTGACTTTGGACT 688

Qy        539 TTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTC 598
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        689 TTAGCTTCAAAGAAAGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTC 748

Qy        599 CAG-AAGCAGAATTGCAAAATTGCTTGGCCCAGGC-TGGCCTGCAGAGTGATACAGAATG 656
              ||| || |||||||||||||||||||||||||||| || ||||||||| ||||||||| ||||
Db        749 CAGAAAACAGAATTGCAAAATTGCTTGGCCCANGCTTGGCCTGCAAAGTGATACAAAATG 808

Qy        657 TACCT----CAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAG 712
              |||||    | || || |  | ||  |  |  || ||||||||||  | ||  |  ||
Db        809 TACCTCCAGCTGTTGATCTTGTGGAGCCCTAAGAAATCATTTTACCTGGAAAATGGTCTC 868

Qy        713 GAATGAAACTGAAAC 727
                |  |||    ||||
Db        869 CAGGGAATGGAAAAC 883
```

RESULT 10
BI259568
LOCUS BI259568 875 bp mRNA linear EST 17-JUL-2001
DEFINITION 602968330F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5108094 5', mRNA sequence.
ACCESSION BI259568
VERSION BI259568.1 GI:14817032
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 875)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11261 row: k column: 07
High quality sequence stop: 722.
FEATURES Location/Qualifiers
source 1. .875
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5108094"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."
ORIGIN

Query Match 47.2%; Score 595; DB 4; Length 875;
Best Local Similarity 91.6%; Pred. No. 5.9e-163;
Matches 809; Conservative 0; Mismatches 50; Indels 24; Gaps 16;

```
Qy        297 CTTTTCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTC 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db          1 CTTTTCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAG-GTTTC 59

Qy        357 GGAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGC 416
              |||||||||||||||||||||||||| |||||||||||||||||||||||||| || ||
Db         60 GGAATTGCTCTTATATTCCAGTCACCA-GGTTGTTCCATGTGGATGATACAGCA-TGTGC 117

Qy        417 CCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCT 476
              |||||||||||  |  || ||||||||||||||||||||||||||||||||||||||||||
Db        118 CCGAGATGTTT--GGAGGCCCTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCT 175

Qy        477 CAAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGG 536
              |||||||||||||| |||||||||||||||||||||||||||||||| ||||
Db        176 CAAACCACGTAACATA-TGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGAC-TTGG 233

Qy        537 ACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGC 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGC 293

Qy        597 TCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATG 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 TCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATG 353

Qy        657 TACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAAT 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 TACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAAT 413

Qy        717 GAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGA 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGA 473

Qy        777 TCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGA 836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 TCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGA 533

Qy        837 CCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATT 896
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        534 CCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCA-T 592
```

```
Qy        897 GTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCT-GGTCATGCTTC 955
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        593 GTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGGTCATGCTTC 652

Qy        956 CCACTCCAGTG--CTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAA-G 1012
              ||||||||| |    |||||||||||||||||||||||||||||||||||||| ||||| |
Db        653 CCACTCCAGGGGCTAAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAACG 712

Qy       1013 AGGAATATGAAG-ATGTTGTAGAAG--ATGTAAAAGAGGAGTGTCAAAAATATGGACCAG 1069
              |||||||||| | |||||||||| |      ||| ||||||||||||  | | |  |
Db        713 AGGAATATGACGAATGTTGTAGACGGATGTAAAAGGAGGAGTGTCAACCAATTTGGACCA 772

Qy       1070 TGGTATCTCTACTTGTT------CCAAAGGAAAATCCTGGCAGAGGACA-AGTCTTTGTT 1122
               |   | |   ||  ||        || | ||||||||||  ||||||| |||||||||
Db        773 GGTGGTATTCTCTACTTTGTTCCAAAAGGAAAAATCCTGGGCGAGGACACAGTCTTTGTA 832

Qy       1123 G-AGTATGCAAAT-GCTGGTGATTCCAAAGCTGCGCAGAAATT 1163
              | | |||||||||| | ||| |||||||||  |||||  || ||
Db        833 GAAATATGCAAATGGTTGGGGATTCCAAAAGTGCGCGAAATTT 875
```

RESULT 11
CF593363

```
LOCUS       CF593363               834 bp    mRNA    linear   EST 26-SEP-2003
DEFINITION  AGENCOURT_15622279 NIH_MGC_147 Homo sapiens cDNA clone
            IMAGE:30531268 5', mRNA sequence.
ACCESSION   CF593363
VERSION     CF593363.1  GI:36346826
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Stefan Hansson
             cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
            and advice from Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM621  row: m  column: 05
            High quality sequence stop: 660.
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30531268"
                     /tissue_type="Human Placenta"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_147"
                     /note="Organ: placenta; Vector: pBluescriptR; Site_1:
                     alI-XhoI; Site_2: BamH; Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTTVN-3', size-selected  for average
                     insert size 2.3 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using  the Cap-trapper method (Carninci, in
                     preparation). Library constructed  by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: This is
                     a NIH_MGC library."
ORIGIN
```

Query Match 47.1%; Score 593.2; DB 7; Length 834;
Best Local Similarity 97.6%; Pred. No. 2e-162;
Matches 645; Conservative 0; Mismatches 9; Indels 7; Gaps 4;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 214

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 274

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 334

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 394

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 454

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 514

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        515 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 574

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        575 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 634

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        635 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTANACTCATTGACTTTGGACTT 694

Qy        541 A-GCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATC-GGGCTC 598
              | |||||||||||| |||||||||||||||||||||||||||||||||||||| ||||||
Db        695 AGGCTTCAAAGAAAGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGGCTC 754

Qy        599 CAGAAGCAGAATTGCAAAATTGCTT---GGCCCAGGCT--GGCCTGCAGAGTGATACAGA 653
              |||||||||||||||||||| |    |   | ||||||||  | ||||||||||||||||
Db        755 CAGAAGCAGAATTGCAAAAATTGCTTGGGCCCCAGGCTTGGCCCTGCAGAGTGATACAGA 814

Qy        654 A 654
              |
Db        815 A 815
```

RESULT 12
BI333581

```
LOCUS       BI333581               682 bp    mRNA    linear   EST 30-JUL-2001
DEFINITION  602997069F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5139184 5',
            mRNA sequence.
ACCESSION   BI333581
VERSION     BI333581.1  GI:15018238
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
```

```
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM11342  row: j  column: 17
           High quality sequence stop: 674.
FEATURES             Location/Qualifiers
     source          1. .682
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5139184"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_12"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
ORIGIN

  Query Match             46.4%;  Score 585.2;  DB 4;  Length 682;
  Best Local Similarity   97.3%;  Pred. No. 4.1e-160;
  Matches  648;  Conservative    0;  Mismatches   13;  Indels    5;  Gaps    5;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 78

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 138

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 198

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 258

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 318

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 378

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        379 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 438

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 GATG-TTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 497

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 557

Qy        541 AGCTTCAAAGAAGGCAATCA-GGATGTAAAGTATATTCAGACAG-ACGGGTATCGGGCTC 598
              |||||||||||||||||||| |||||||||||||||||||||| ||||||||||||||||
Db        558 AGCTTCAAAGAAGGCAATCACGGATGTAAAGTATATTCAGACAGAACGGGTATCGGGCTC 617

Qy        599 CAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATA-CAGAATGT 657
              ||  |      | |   ||| ||||||| |||||||||||||||||||| ||||||||
Db        618 CAAGAAGCAGAATTGCAAAATGCTTGG-CCAGGCTGGCCTGCAGAGTGATACCAGAATGT 676

Qy        658 ACCTCA 663
              ||||||
Db        677 ACCTCA 682
```

```
RESULT 13
BG386905
LOCUS       BG386905                906 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  602454704F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4583065 5',
            mRNA sequence.
ACCESSION   BG386905
VERSION     BG386905.1  GI:13280454
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: NIH Intramural Sequencing Center
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM1306  row: g  column: 02
           High quality sequence stop: 668.
FEATURES             Location/Qualifiers
     source          1. .906
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4583065"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_15"
                     /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)"
ORIGIN

  Query Match             46.3%;  Score 583.8;  DB 4;  Length 906;
  Best Local Similarity   96.1%;  Pred. No. 1.2e-159;
  Matches  664;  Conservative    0;  Mismatches   17;  Indels   10;  Gaps    6;

Qy        570 GTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCA 629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCA 61

Qy        630 GGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAAT 689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAAT 121

Qy        690 CATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATG 749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATG 181

Qy        750 GAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGC 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGC 241

Qy        810 CGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAG 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAG 301

Qy        870 CAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCC 929
```

```
Db        302 CAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCC 361

Qy        930 TCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGA 989

Db        362 TCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGA 421

Qy        990 TGATGATTATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGAGGA 1049

Db        422 TGATGATTATCTTGAGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGAGGA 481

Qy       1050 GTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGG 1109

Db        482 GTGTCAAAAATATGGACCAGTGGTATCTCTA-TTGTTCCAAAGGAAAATCCTGG-AGAGG 539

Qy       1110 ACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATTACTGAC 1169

Db        540 ACAAGTCTTTGTTGAGTATGCAAATG-TGGTGATTCCAAAGCTGCGCAGAAATTACTGAC 598

Qy       1170 TGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGCCTACAA 1229

Db        599 TGGAAG--ATGTTGATGGGAAGTTGTGT---GGTACATTCTACCCCGTGAGTGC--TACA 651

Qy       1230 GAGGGGATATCTGTATCAAACCTTGCTTTAA 1260

Db        652 AGAGGGATATCTGTATCAAACCTGGCTTTAA 682


RESULT 14
BQ225071
LOCUS       BQ225071                888 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7323233 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6016796
            5', mRNA sequence.
ACCESSION   BQ225071
VERSION     BQ225071.1  GI:20406471
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 888)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13215  row: a  column: 21
            High quality sequence stop: 581.
FEATURES             Location/Qualifiers
     source          1. .888
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6016796"
                     /tissue_type="epithelioid carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_70"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.1 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             46.0%;  Score 579.6;  DB 5;  Length 888;
  Best Local Similarity   97.9%;  Pred. No. 2e-158;
  Matches  619;  Conservative    0;  Mismatches    9;  Indels    4;  Gaps    3;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Db        149 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 208

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Db        209 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 268

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Db        269 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 328

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Db        329 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 388

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Db        389 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 448

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Db        449 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 508

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420

Db        509 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 568

Qy        421 GATGTTTTGGAGG-CCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAA 479

Db        569 GATGTTTTGGAGGCCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAA 628

Qy        480 ACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACT 539

Db        629 ACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACT 688

Qy        540 TAGCTTCAAAGAA-GGCAATCAGGATGTAAAGTATATTCAGACAGACGGGT--ATCGGGC 596

Db        689 TAGCTTCAAAGAAGGGAAATCAAGATGTAAAGTATATTCAGACCGACGGGTTATCGGGGC 748

Qy        597 TCCAGAAGCAGAATTGCAAAATTGCTTGGCCC 628

Db        749 TCCAGAAGCAAAATTGCAAAATTGCTTGGGCC 780


RESULT 15
CO737403
LOCUS       CO737403                602 bp    mRNA    linear   EST 29-JUL-2004
DEFINITION  S1LT02c17i16f1 squirrel testis library 1 Spermophilus lateralis
            cDNA clone 17i16 5', mRNA sequence.
ACCESSION   CO737403
VERSION     CO737403.1  GI:50824673
KEYWORDS    EST.
SOURCE      Spermophilus lateralis (golden-mantled ground squirrel)
  ORGANISM  Spermophilus lateralis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
            Spermophilus.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
            Rogers,J. and Cossins,A.R.
  TITLE     Microarray analysis of transcriptional changes during hibernation
            in the golden mantled ground squirrel, Spermophilus lateralis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Andrew R. Cossins
            Laboratory for Environmental Gene Regulation
            University of Liverpool
            School of Biological Sciences, The Biosciences Building, Crown
            Street, Liverpool, United Kingdom, L69 7ZB
            Tel: +44(0)151-795-4510
            Fax: +44(0)151-795-4431
            Email: cossins@liv.ac.uk
```

```
            Vector has been trimmed from this EST.
            Plate: 17  row: i  column: 16
            Seq primer: pflc T7 (5'-AATACGACTCACTATAGGG-3')
            High quality sequence stop: 602.
FEATURES              Location/Qualifiers
     source           1. .602
                      /organism="Spermophilus lateralis"
                      /mol_type="mRNA"
                      /db_xref="taxon:76772"
                      /clone="17i16"
                      /sex="Male & female"
                      /tissue_type="Testis"
                      /dev_stage="Adult"
                      /lab_host="E.coli Electromax DH10B"
                      /clone_lib="squirrel testis library 1"
                      /note="Vector: pPLC; Site_1: SalI GTCGAG; Site_2: BamHI
                      GGATCC; Normalized and subtracted cDNA library prepared
                      from testis of hibernating and summer animals"
ORIGIN

  Query Match             44.4%;  Score 560;  DB 7;  Length 602;
  Best Local Similarity   95.8%;  Pred. No. 1e-152;
  Matches  575;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy        624 GGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCT 62

Qy        684 AGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCA 743
              |||||| |||||||||||||||||||||||||||||||||||||||| ||||||||
Db         63 AGGAATTATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTTAGATCTCA 122

Qy        744 GGAATGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGT 803
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        123 GGAATGGAAGGCAAATAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGT 182

Qy        804 GAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 242

Qy        864 TCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTT 923
              |||||||||||| |||||||||||||||||||||||| |||||||| ||||||||||||
Db        243 TCCAAGCAGAAGGATTCCTGCTGAAATGGCATTGTGCAGTCCATTCTTCAGCATTCCTTT 302

Qy        924 TGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGT 983
              |||||||||||||||||| |||| |||||||| |||||||| |||||| |||| |||||
Db        303 TGCCCCTCATATTGAAGATTTGGTGATGCTTCCTACTCCAGTCCTAAGATTGCTCAATGT 362

Qy        984 GCTGGATGATGATTATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAA 1043
              ||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||
Db        363 CCTGGATGATGATTATCTTGAGAATGAAGAAGAATATGAAGATGTTGTAGAAGATGTAAA 422

Qy       1044 AGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGG 1103
              |||||| |||||||||||||||||| || ||||| ||||||||||||||||||||||||
Db        423 AGAGGAATGTCAAAAATATGGACCAGTAGTGTCTCTGCTTGTTCCAAAGGAAAATCCTGG 482

Qy       1104 CAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATT 1163
              |||||| |||||||||||||||||||||| |||||||||||||||||| ||||||||
Db        483 CAGAGGGCAAGTCTTTGTTGAGTATGCAAATCCTGGTGATTCCAAAGCTGCTCAGAAATT 542

Qy       1164 ACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGC 1223
              ||| |||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db        543 ACTAACTGGAAGGATGTTTGATGGGAAATTCGTTGTGGCTACATTCTACCCGCTGAGTGC 602


Search completed: November 29, 2004, 17:46:23
Job time : 5316 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:30 ; Search time 5510 Seconds
(without alignments)
10813.978 Million cell updates/sec

Title: US-10-798-532-1
Perfect score: 1260
Sequence: 1 atggcgggatccggctgcgc.....tgtatcaaaccttgctttaa 1260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1256.8	99.7	1260	6	CQ715025	CQ715025 Sequence
2	1256.8	99.7	1260	6	AX262521	AX262521 Sequence
3	1256.8	99.7	2598	6	AX262519	AX262519 Sequence
4	1256.8	99.7	2909	9	HSA536197	AJ536197 Homo sapi
5	1256.8	99.7	4065	6	CQ412057	CQ412057 Sequence
6	1256.8	99.7	4065	6	AX188296	AX188296 Sequence
7	1255.2	99.6	1260	6	AX166543	AX166543 Sequence
8	1255.2	99.6	1824	6	AX680136	AX680136 Sequence
9	1116	88.6	2041	10	BC058732	BC058732 Mus muscu
c 10	1112.8	88.3	110000	2	AC129141_1	Continuation (2 of
11	1108	87.9	1633	10	RNKISRNA	X98374 R.norvegicu
12	1108	87.9	3244	10	RNU70372	U70372 Rattus norv
13	1100	87.3	1254	10	MMPKKIS	Y10725 Mus musculu
14	1067.2	84.7	1367	9	BC026046	BC026046 Homo sapi
15	989.8	78.6	1135	9	AK058195	AK058195 Homo sapi
16	941.8	74.7	2041	9	BC014917	BC014917 Homo sapi
17	769.6	61.1	1157	5	CR386837	CR386837 Gallus ga
18	699.8	55.5	2940	9	HSM805327	AL834136 Homo sapi
c 19	567.6	45.0	259562	2	AC094499	AC094499 Rattus no
20	535.4	42.5	601	10	MMKISRNA	X82320 M.musculus
21	458.6	36.4	529	6	AX186343	AX186343 Sequence
22	445.8	35.4	722	6	BD221413	BD221413 Human gen
23	426.6	33.9	470	6	AX187546	AX187546 Sequence
24	387.4	30.7	462	6	AX184680	AX184680 Sequence
25	295	23.4	105902	9	AL359699	AL359699 Human DNA
26	295	23.4	167995	2	AC040901	AC040901 Homo sapi
c 27	293.4	23.3	211812	2	AC027205	AC027205 Homo sapi
c 28	254.4	20.2	200095	10	AC121606	AC121606 Mus muscu
29	254.4	20.2	220806	10	AC123650	AC123650 Mus muscu
30	252.8	20.1	100207	10	AY180177	AY180177 Mus muscu
31	249.6	19.8	211635	2	AC112763	AC112763 Rattus no
32	203.8	16.2	231	6	BD220137	BD220137 Human gen
33	186.4	14.8	185810	2	AC136052	AC136052 Rattus no
34	179	14.2	419	6	CQ393016	CQ393016 Sequence
35	179	14.2	419	6	CQ399411	CQ399411 Sequence
36	166.2	13.2	60640	10	AY180176	AY180176 Mus muscu
37	155	12.3	205752	2	AC119431	AC119431 Mus muscu
c 38	87.4	6.9	185810	2	AC136052	AC136052 Rattus no
39	85	6.7	198	6	CQ405796	CQ405796 Sequence
40	69.6	5.5	2108	8	NPL18351	Y18351 Nicotiana p
41	64.6	5.1	2047	8	BT013617	BT013617 Lycopersi
42	63.4	5.0	65	6	CQ555763	CQ555763 Sequence
43	60.6	4.8	1996	5	BC044032	BC044032 Xenopus l
44	59.8	4.7	2204	8	NPL18350	Y18350 Nicotiana p
45	55.8	4.4	1565	5	BC065869	BC065869 Danio rer

ALIGNMENTS

RESULT 1
CQ715025
LOCUS CQ715025 1260 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 959 from Patent WO02068579.
ACCESSION CQ715025
VERSION CQ715025.1 GI:42275882
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 959 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1. .1260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 99.7%; Score 1256.8; DB 6; Length 1260;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
```

```
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

RESULT 2

```
AX262521
LOCUS       AX262521                1260 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 12 from Patent WO0173050.
ACCESSION   AX262521
VERSION     AX262521.1  GI:16511408
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Meyers,R.
  TITLE     3714, 16742, 23546, and 13887 novel protein kinase molecules and
            uses therefor
  JOURNAL   Patent: WO 0173050-A 12 04-OCT-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1260
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             99.7%;  Score 1256.8;  DB 6;  Length 1260;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1258;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

```
RESULT 3
AX262519
LOCUS       AX262519                2598 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 10 from Patent WO0173050.
ACCESSION   AX262519
VERSION     AX262519.1  GI:16511406
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Meyers,R.
  TITLE     3714, 16742, 23546, and 13887 novel protein kinase molecules and
            uses therefor
  JOURNAL   Patent: WO 0173050-A 10 04-OCT-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2598
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             268. .1527
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD10238.1"
                     /db_xref="GI:16511407"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
```

```
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

  Query Match           99.7%;  Score 1256.8;  DB 6;  Length 2598;
  Best Local Similarity 99.8%;  Pred. No. 0;
  Matches 1258;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1527


RESULT 4
HSA536197
LOCUS       HSA536197               2909 bp    mRNA    linear   PRI 11-JAN-2003
DEFINITION  Homo sapiens mRNA for KIS protein.
ACCESSION   AJ536197
VERSION     AJ536197.1  GI:27657360
KEYWORDS    KIS protein.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Bieche,I., Manceau,V., Curmi,P.A., Laurendon,I., Lachkar,S.,
            Leroy,K., Vidaud,D., Sobel,A. and Maucuer,A.
  TITLE     Quantitative RT-PCR reveals a ubiquitous but preferentially neural
            expression of the KIS gene in rat and human
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2909)
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-2003) Maucuer A., U440, Inserm, IFM, 17 rue du
            fer a moulin, 75005 Paris, FRANCE
FEATURES             Location/Qualifiers
     source          1. .2909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:IMAGp998M213680"
                     /db_xref="taxon:9606"
                     /clone="I.M.A.G.E. Consortium CloneID 1455548"
                     /tissue_type="2 pooled tumors (clear cell type)"
     gene            1. .2909
                     /gene="KIS"
     CDS             146. .1405
                     /gene="KIS"
                     /codon_start=1
                     /product="KIS protein"
                     /protein_id="CAD60192.1"
                     /db_xref="GI:27657361"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

  Query Match             99.7%;  Score 1256.8;  DB 9;  Length 2909;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1258;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        146 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 205

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 265

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 325

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 385

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 445

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 505

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        506 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 565

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 625

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 685

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 745

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 805

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 865

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 925

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 985

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1045

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1105

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1106 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1165

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1225

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1285
```

```
Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1345

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1346 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1405


RESULT 5
CQ412057
LOCUS       CQ412057                4065 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 19128 from Patent WO0170979.
ACCESSION   CQ412057
VERSION     CQ412057.1  GI:41319838
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Lillie,J.
  TITLE     Genes, compositions, kits, and method for identification,
            assessment, prevention, and therapy of ovarian cancer
  JOURNAL   Patent: WO 0170979-A 19128 27-SEP-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .4065
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             99.7%;  Score 1256.8;  DB 6;  Length 4065;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1258;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1527


RESULT 6
AX188296
LOCUS       AX188296                4065 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 3991 from Patent WO0142467.
ACCESSION   AX188296
VERSION     AX188296.1  GI:15139769
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
  TITLE     Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
  JOURNAL   Patent: WO 0142467-A 3991 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .4065
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             99.7%;  Score 1256.8;  DB 6;  Length 4065;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1258;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1527
```

RESULT 7
AX166543

```
LOCUS       AX166543               1260 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 34 from Patent WO0138503.
ACCESSION   AX166543
VERSION     AX166543.1  GI:14546888
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
            Flanagan,P. and Clary,D.S.
  TITLE     Novel human protein kinases and protein kinase-like enzymes
  JOURNAL   Patent: WO 0138503-A 34 31-MAY-2001;
            Sugen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1260
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             99.6%;  Score 1255.2;  DB 6;  Length 1260;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1257;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420
```

```
Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

```
RESULT 8
AX680136
LOCUS       AX680136                1824 bp    DNA     linear   PAT 29-MAR-2003
DEFINITION  Sequence 21 from Patent WO0181555.
ACCESSION   AX680136
VERSION     AX680136.1  GI:29369916
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Gururajan,R., Patterson,C., Recipon,S.A., Tribouley,C.M.,
            Hafalia,A., Khan,F.Y., Yue,H., Au-Young,J., Bandman,O.,
            Baughn,M.R., Borowsky,M.L., Burford,N., Burrill,J.D., Elliot,V.S.,
            Gandhi,A.R., Kearney,L., Lal,P., Lu,D.A., Lu,Y., Tang,Y.T.,
            Zingler,K.A., Griffin,J.A., Hillman,J.J., Marcus,G.A., Nguyen,D.B.,
            Policky,J., Ramkumar,J., Thornton,M., Walia,N.K. and Walsh,R.T.
  TITLE     Human kinases
  JOURNAL   Patent: WO 0181555-A 21 01-NOV-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1824
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 3107952CB1"
ORIGIN
```

```
  Query Match             99.6%;  Score 1255.2;  DB 6;  Length 1824;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1257;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 228

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 288

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 348

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 408

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 468

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 528

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        529 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 588

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 648

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 708

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 768

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 828

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 888

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 948

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1008
```

```
Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1068

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1128

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1129 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1188

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1189 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1248

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1249 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1308

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1309 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1368

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1369 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1428
```

RESULT 9
BC058732
LOCUS BC058732 2041 bp mRNA linear ROD 30-JUN-2004
DEFINITION Mus musculus kinase interacting with leukemia-associated gene (stathmin), mRNA (cDNA clone MGC:64770 IMAGE:6414877), complete cds.
ACCESSION BC058732
VERSION BC058732.1 GI:37194892
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2041)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 2041)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
Web site: http://genome.uiowa.edu
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A., Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K., Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J., Casavant,T., Soares,M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754447.

FEATURES Location/Qualifiers
source 1. .2041
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:64770 IMAGE:6414877"
/tissue_type="Brain, pool of enriched mouse brain tissues from 13.5,14.5,16.5 and 17.5 dpc"
/clone_lib="NIH_BMAP_FR0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
gene 1. .2041
/gene="Kist"
/note="synonym: KIS"
/db_xref="LocusID:16589"
/db_xref="MGI:1341908"
CDS 154. .1413
/gene="Kist"
/codon_start=1
/product="kinase interacting with leukemia-associated gene (stathmin)"
/protein_id="AAH58732.1"
/db_xref="GI:37194893"
/db_xref="LocusID:16589"
/db_xref="MGI:1341908"
/translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
MFDGKFVVATFYPLSAYKRGYLYQTLL"

ORIGIN

Query Match 88.6%; Score 1116; DB 10; Length 2041;
Best Local Similarity 92.9%; Pred. No. 1.5e-278;
Matches 1170; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              |||||||| |||||||||||| |||||||| |||||||||||  || ||||||||||||||
Db        154 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 213

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              |||||||||||||| |||||||| ||||| |||||||| || |||||||||||| |||||
Db        214 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 273

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              || |||||||| | ||| |||||||||||||||||||||||||||||| |||| || |||
Db        274 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 333

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||
```

```
Db        334 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 393
Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
Db        394 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 453
Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
Db        454 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 513
Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
Db        514 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 573
Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
Db        574 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 633
Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
Db        634 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 693
Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
Db        694 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 753
Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
Db        754 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 813
Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
Db        814 TCAGCTGTTGATCTGTGGAGCCTCGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAG 873
Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
Db        874 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAT 933
Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
Db        934 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 993
Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
Db        994 ATCAAAAGCATGCTTCATGATGACCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGC 1053
Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
Db       1054 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACT 1113
Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
Db       1114 CCAGTGCTCAGGCTCCTCAATGTGCTGGATGATGATTACCTTGAAAATGAAGATGAATAT 1173
Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
Db       1174 GAAGATGTTGTAGAAGATGTGAAAGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1233
Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
Db       1234 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1293
Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
Db       1294 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1353
Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
Db       1354 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1413

RESULT 10
AC129141_1/c
```

```
WPCOMMENT
  Sequence split into 4 fragments  LOCUS AC129141 Accession AC129141
     Fragment Name      Begin     End
     AC129141_0             1    110000
     AC129141_1        100001    210000
     AC129141_2        200001    310000
     AC129141_3        300001    363206
Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH23

  Query Match           88.3%;  Score 1112.8;  DB 2;  Length 110000;
  Best Local Similarity   92.7%;  Pred. No. 2.3e-277;
  Matches 1168;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
Db      70316 ATGGCCGGGTCCTGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 70257
Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
Db      70256 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 70197
Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
Db      70196 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 70137
Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
Db      70136 ACCACCGGGGCTGCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 70077
Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
Db      70076 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 70017
Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
Db      70016 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 69957
Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
Db      69956 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATACAGCACTGTGCCAGA 69897
Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
Db      69896 GATGTTCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAA 69837
Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
Db      69836 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 69777
Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
Db      69776 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 69717
Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
Db      69716 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 69657
Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
Db      69656 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 69597
Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
Db      69596 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 69537
Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
Db      69536 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 69477
Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
Db      69476 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 69417
```

```
Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||
Db      69416 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 69357

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||| ||||| || ||||||||||||||||| |||||||  |||||||| ||||||
Db      69356 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 69297

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              |||||||||||||||||||||||||| |||||||| ||||||||||||||||| ||||| 
Db      69296 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 69237

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||| 
Db      69236 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 69177

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              |||||||||||||| |||||||| |||||||| |||||||||||||||||||||||||||
Db      69176 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGATGGGAAGTTTGTTGTG 69117

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      69116 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 69057
```

```
RESULT 11
RNKISRNA
LOCUS       RNKISRNA                1633 bp    mRNA    linear   ROD 08-JUL-2002
DEFINITION  R.norvegicus mRNA for KIS protein.
ACCESSION   X98374
VERSION     X98374.1  GI:1403531
KEYWORDS    kis gene.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Maucuer,A., Ozon,S., Manceau,V., Gavet,O., Lawler,S., Curmi,P. and
            Sobel,A.
  TITLE     KIS is a protein kinase with an RNA recognition motif
  JOURNAL   J. Biol. Chem. 272 (37), 23151-23156 (1997)
  MEDLINE   97435279
   PUBMED   9287318
REFERENCE   2  (bases 1 to 1633)
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-1996) A. Maucuer, INSERM U440, 17 rue du Fer
            Moulin, 75005 Paris, FRANCE
FEATURES             Location/Qualifiers
     source          1..1633
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /tissue_type="brain"
                     /clone_lib="7"
                     /dev_stage="neonate"
     gene            1..1633
                     /gene="KIS"
     CDS             39..1298
                     /gene="KIS"
                     /codon_start=1
                     /protein_id="CAA67021.1"
                     /db_xref="GI:1403532"
                     /db_xref="GOA:Q63285"
                     /db_xref="Swiss-Prot:Q63285"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

  Query Match             87.9%;  Score 1108;  DB 10;  Length 1633;
  Best Local Similarity   92.5%;  Pred. No. 1.7e-276;
  Matches 1165;  Conservative    0;  Mismatches   95;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||| || ||||||||||| |||||||| ||||||||||| || |||||||||||||||
Db         39 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 98

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||| || || ||||||||||| ||||||||||| ||||| 
Db         99 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 158

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||  || |||||||||||||||||||||||||||||| |||| || |||
Db        159 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 218

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||| ||||| || || |||||||||||||| ||||||||||| ||||||
Db        219 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 278

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              || ||||||||||||||||| |||||||||||||| || ||||| ||||| || ||||| 
Db        279 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 338

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        339 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 398

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||  |||||||||||||||||||||| |||||||||||||| ||||| || ||| ||
Db        399 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 458

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||  |||||||||||||||||||||| |||||||||||||||||||| |||||||| 
Db        459 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 518

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||| |||||  |||||||||| ||||| || || ||||| || |||||||||||||| 
Db        519 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 578

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||| || |||||||||||||||||||||||||| ||||| 
Db        579 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 638

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||| |||| ||||||||||||||||| |||||||||||||||||||| ||||||
Db        639 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 698

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db        699 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 758

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||| 
Db        759 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 818

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        819 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 878

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              |||||||||||||||||||| || ||||||||||| || |||||||| ||||| ||||||
Db        879 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 938

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||
Db        939 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 998

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||| ||||| || ||||||||||||||||| |||||||  |||||||| ||||||
```

```
Db        999 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1058

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              |||||||||||||||||||||||| |||||||| |||||||||||||||| |||||
Db       1059 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1118

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db       1119 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1178

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              |||||||||||||| |||||||| |||||||| |||||||||| ||||||||||||||||
Db       1179 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1238

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1239 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1298


RESULT 12
RNU70372
LOCUS       RNU70372                3244 bp    mRNA    linear   ROD 08-SEP-1999
DEFINITION  Rattus norvegicus PAM COOH-terminal interactor protein 2 mRNA,
            complete cds.
ACCESSION   U70372
VERSION     U70372.2  GI:5821767
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 3244)
  AUTHORS   Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
            and Eipper,B.A.
  TITLE     Novel proteins that interact with the COOH-terminal cytosolic
            routing determinants of an integral membrane peptide-processing
            enzyme
  JOURNAL   J. Biol. Chem. 271 (45), 28636-28640 (1996)
  MEDLINE   97067094
   PUBMED   8910496
REFERENCE   2  (bases 1 to 3244)
  AUTHORS   Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
            Eipper,B.A. and Mains,R.E.
  TITLE     The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
            Determinants of the Peptide Processing Enzyme Peptidylglycine
            alpha-Amidating Monooxygenase
  JOURNAL   J. Biol. Chem. 274 (1999) In press
REFERENCE   3  (bases 1 to 3244)
  AUTHORS   Alam,R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
            and Eipper,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
            School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
            21205, USA
REFERENCE   4  (bases 1 to 3244)
  AUTHORS   Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
            Eipper,B.A. and Mains,R.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
            School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
            21205, USA
  REMARK    Sequence update by submitter
COMMENT     On Sep 3, 1999 this sequence version replaced gi:1698778.
FEATURES             Location/Qualifiers
     source          1. .3244
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
     CDS             162. .1421
                     /function="serine protein kinase"
                     /note="P-CIP2"
                     /codon_start=1
                     /product="PAM COOH-terminal interactor protein 2"
                     /protein_id="AAC53031.2"
                     /db_xref="GI:5821768"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

  Query Match           87.9%;  Score 1108;  DB 10;  Length 3244;
  Best Local Similarity  92.5%;  Pred. No. 2e-276;
  Matches 1165;  Conservative    0;  Mismatches   95;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||| || |||||||||||| |||||||| |||||||||| || ||||||||||||||||
Db        162 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 221

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              |||||||||||||||||||||||||||| || || |||||||||||| |||||||||| |||||
Db        222 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 281

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||  || |||||||||||||||||||||||||||| |||| || |||
Db        282 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 341

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||| ||||| || || |||||||||||||| |||||||||| ||||||
Db        342 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 401

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              || |||||||||||||||| |||||||||||||| || ||||| ||||| || |||||
Db        402 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 461

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        462 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 521

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||  |||||||||||||||||||||| |||||||||||||| ||||| || ||| ||
Db        522 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 581

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||  |||||||||||||||||||||||||||||||||||||||| |||||||||
Db        582 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 641

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||| |||||  |||||||||| ||||| || || ||||| || ||||||||||||||
Db        642 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 701

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||||| || |||||||||||||||||||||||||| ||||||
Db        702 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 761

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||| |||| |||||||||||||||| |||||||||||||||||| ||||||
Db        762 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 821

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db        822 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 881

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        882 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 941

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        942 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 1001
```

```
Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||| || |||||||||| || ||||||||| ||||| ||||||
Db       1002 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 1061

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||
Db       1062 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 1121

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||| ||||| || |||||||||||||||| |||||| |||||||| ||||||
Db       1122 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1181

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              |||||||||||||||||||||||| |||||||| |||||||||||||||| |||||
Db       1182 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1241

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db       1242 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1301

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              |||||||||||||| |||||||| |||||||||| |||||||||||||||| ||||||
Db       1302 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1361

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1362 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1421


RESULT 13
MMPKKIS
LOCUS       MMPKKIS                 1254 bp    mRNA    linear   ROD 08-JUL-2002
DEFINITION  Mus musculus partial mRNA for KIS protein kinase (Kis gene).
ACCESSION   Y10725
VERSION     Y10725.2  GI:21726712
KEYWORDS    Kis gene; protein kinase.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1254)
  AUTHORS   Maucuer,A., Ozon,S., Manceau,V., Gavet,O., Lawler,S., Curmi,P. and
            Sobel,A.
  TITLE     KIS is a protein kinase with an RNA recognition motif
  JOURNAL   J. Biol. Chem. 272 (37), 23151-23156 (1997)
  MEDLINE   97435279
   PUBMED   9287318
REFERENCE   2
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-1997) Maucuer A., INSERM U440, U440, 17 Rue du
            Fer Moulin, Paris, 75005, FRANCE
  REMARK    revised by [4]
REFERENCE   3  (bases 1 to 1254)
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) A. Maucuer, INSERM U440, U440, 17 Rue du
            Fer Moulin, Paris, 75005, FRANCE
COMMENT     On Jul 10, 2002 this sequence version replaced gi:1806129.
            Overlaps with related sequences X82320, X10725.
FEATURES             Location/Qualifiers
     source          1. .1254
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="K22"
                     /tissue_type="brain"
                     /clone_lib="lambda gt10"
                     /dev_stage="neonate"
     gene            1. .1254
                     /gene="Kis"
     CDS             11. .>1254
                     /gene="Kis"
                     /codon_start=1
                     /product="KIS protein kinase"
                     /protein_id="CAA71714.2"
                     /db_xref="GI:21726713"
                     /db_xref="GOA:P97343"
                     /db_xref="Swiss-Prot:P97343"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYL"
ORIGIN

  Query Match           87.3%;  Score 1100;  DB 10;  Length 1254;
  Best Local Similarity   92.8%;  Pred. No. 1.9e-274;
  Matches 1154;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              |||||||| |||||||||| |||||||| |||||||||| || ||||||||||||||||
Db         11 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 70

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              |||||||||||||| |||||||| ||||| |||||||| || |||||||||| |||||
Db         71 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 130

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              || |||||||| | ||| |||||||||||||||||||||||||||||| |||| || |||
Db        131 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 190

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||||
Db        191 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 250

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              || |||||||||||||||| |||||||||||| |||||||| ||||| || |||||
Db        251 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 310

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        311 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 370

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||| |||||||||||||| ||||| || |||||||||||||||||| || ||||||
Db        371 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 430

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||| |||||||||| |||||||||| |||||||||| |||||||| |||||||||
Db        431 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 490

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||| |||||  |||||||||| |||||||| |||||||| || ||||||||||||||
Db        491 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 550

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||| |||||
Db        551 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 610

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              |||||||  |||| || |||||||||||||| ||||||||||||||||||||||||||||
Db        611 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 670

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        671 TCAGCTGTTGATCTGTGGAGCCTCGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAG 730

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        731 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAT 790

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
```

```
Db        791 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 850

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||  ||  ||||||||||  ||||||||  ||||||||||||
Db        851 ATCAAAAGCATGCTTCATGATGACCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGC 910

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||  |||
Db        911 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACT 970

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              ||||||||  ||  ||  ||  ||||||||||||||||||||||  ||||  ||||||||  ||||||
Db        971 CCAGTGCTCAGGCTCCTCAATGTGCTGGATGATGATTACCTTGAAAATGAAGATGAATAT 1030

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||  ||||||||||||||  ||||||||||||||||||  |||||
Db       1031 GAAGATGTTGTAGAAGATGTGAAAGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1090

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||||||||||||||||||||||||  ||||||||  ||  ||  ||||||
Db       1091 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1150

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||  |||||  ||  ||||||||  ||||||||||  ||||||||||||||||
Db       1151 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1210

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTA 1244
              |||||||||||||||||||||||||||||||||||||||||  ||
Db       1211 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTA 1254
```

```
RESULT 14
BC026046
LOCUS       BC026046                1367 bp    mRNA    linear   PRI 30-JUN-2004
DEFINITION  Homo sapiens kinase interacting with leukemia-associated gene
            (stathmin), mRNA (cDNA clone MGC:39391 IMAGE:4649261), complete
            cds.
ACCESSION   BC026046
VERSION     BC026046.1  GI:19684094
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1367)
  AUTHORS   Strausberg,R.L., Peingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1367)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 42 Row: f Column: 6
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 21389438.
FEATURES             Location/Qualifiers
     source          1. .1367
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:39391 IMAGE:4649261"
                     /tissue_type="Eye, retinoblastoma"
                     /clone_lib="NIH_MGC_16"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1. .1367
                     /gene="KIS"
                     /db_xref="LocusID:127933"
     CDS             159. .1193
                     /gene="KIS"
                     /codon_start=1
                     /product="kinase interacting stathmin"
                     /protein_id="AAH26046.1"
                     /db_xref="GI:19684095"
                     /db_xref="LocusID:127933"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSATIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEGLC"
ORIGIN
```

```
  Query Match             84.7%;  Score 1067.2;  DB 9;  Length 1367;
  Best Local Similarity   92.7%;  Pred. No. 6.5e-266;
  Matches 1168;  Conservative    0;  Mismatches    3;  Indels   89;  Gaps    1;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 218

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 278

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 338

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 398
```

```
Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 458

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 518

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        519 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 578

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 638

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 698

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 758

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 818

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 878

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        879 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTACTATTGATCAC 938

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 998

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1058

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1118

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1119 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1178

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||
Db       1179 GAAG---------------------------------------------------------- 1182

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
                                                 ||||||||||||||||||||||||||||
Db       1183 ---------------------------------GTCTTTGTTGAGTATGCAAATGCTGGT 1209

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1269

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1270 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1329
```

RESULT 15
AK058195
LOCUS AK058195 1135 bp mRNA linear PRI 12-SEP-2003

DEFINITION Homo sapiens cDNA FLJ25466 fis, clone TST09304, highly similar to Kinase interacting with leukemia-associated gene (stathmin).
ACCESSION AK058195
VERSION AK058195.1 GI:16554273
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T., Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R., Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiyama,H., Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E., Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S., Terashima,Y., Watanabe,M., Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1135)
AUTHORS Sugano,S. and Suzuki,Y.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.
FEATURES Location/Qualifiers
source 1. .1135
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TST09304"
/tissue_type="testis"
/clone_lib="TST"
/note="cloning vector: pME18SFL3"
ORIGIN

Query Match 78.6%; Score 989.8; DB 9; Length 1135;
Best Local Similarity 99.8%; Pred. No. 8e-246;
Matches 991; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        268 GTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTG 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTG 182

Qy        328 TTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGT 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 TTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGT 242

Qy        388 TGTTCCATGTGGATGATACAGCATTGCGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTT 447
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        243 TGTTCCATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTT 302

Qy        448 CATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAG 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAG 362

Qy        508 AATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTA 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTA 422

Qy        568 AAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCC 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        423 AAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCC 482

Qy        628 CAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGA 542

Qy        688 ATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAA 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 ATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAA 602

Qy        748 TGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAAT 807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 TGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAAT 662

Qy        808 GCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCA 867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 GCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCA 722

Qy        868 AGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCC 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 AGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCC 782

Qy        928 CCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTG 987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 CCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTG 842

Qy        988 GATGATGATTATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGAG 1047
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        843 GATGATGATTATCTTGAGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGAG 902

Qy       1048 GAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGA 1107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGA 962

Qy       1108 GGACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATTACTG 1167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GGACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATTACTG 1022

Qy       1168 ACTGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGCCTAC 1227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 ACTGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGCCTAC 1082

Qy       1228 AAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||||||||||||||||
Db       1083 AAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1115
```

Search completed: November 29, 2004, 14:49:16
Job time : 5517 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:25 ; Search time 653.5 Seconds
(without alignments)
10121.301 Million cell updates/sec

Title: US-10-798-532-1
Perfect score: 1260
Sequence: 1 atggcgggatccggctgcgc.....tgtatcaaaccttgctttaa 1260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1260	100.0	1260	3	AAZ51355	Aaz51355 Human wil
	2	1258.4	99.9	1260	3	AAZ51356	Aaz51356 Human tra
	3	1258.4	99.9	2008	6	ABQ99383	Abq99383 Human cod
	4	1256.8	99.7	1260	4	AAI66829	Aai66829 Human pro
	5	1256.8	99.7	2622	4	AAI66828	Aai66828 Human pro
	6	1256.8	99.7	4065	4	AAH72714	Aah72714 Human cer
	7	1256.8	99.7	4065	5	ADL45238	Adl45238 Human ova
	8	1255.2	99.6	1260	4	AAS06734	Aas06734 Polynucle
	9	1255.2	99.6	1824	4	AAD18818	Aad18818 Human kin
	10	1175	93.3	3998	12	ADQ24401	Adq24401 Human sof
	11	1108	87.9	3244	10	ADB52857	Adb52857 Primary r
	12	458.6	36.4	529	4	AAH70764	Aah70764 Human cer
	13	445.8	35.4	722	3	AAA02535	Aaa02535 Human col
	14	426.6	33.9	470	4	AAH71966	Aah71966 Human cer
	15	375.4	29.8	461	4	AAH69101	Aah69101 Human cer
	16	272.2	21.6	411	6	ABQ99225	Abq99225 Human ORF
	17	203.8	16.2	231	3	AAA01259	Aaa01259 Human col
	18	179	14.2	419	5	ADI73740	Adi73740 Human ova
	19	179	14.2	419	5	ADI67345	Adi67345 Human ova
c	20	177.8	14.1	546	6	ABQ53127	Abq53127 Oligonucl
	21	177.8	14.1	546	6	ABQ53126	Abq53126 Oligonucl
c	22	177.6	14.1	547	6	ABQ44465	Abq44465 Oligonucl
	23	177.6	14.1	547	6	ABQ44464	Abq44464 Oligonucl
	24	158.8	12.6	546	6	ABQ53129	Abq53129 Oligonucl
c	25	158.8	12.6	546	6	ABQ53128	Abq53128 Oligonucl
	26	158.8	12.6	547	6	ABQ44463	Abq44463 Oligonucl
c	27	158.8	12.6	547	6	ABQ44462	Abq44462 Oligonucl
	28	148.4	11.8	442	9	ACH28557	Ach28557 Human adu
	29	85	6.7	198	5	ADL38977	Adl38977 Human ova
	30	63.4	5.0	65	6	ABN52650	Abn52650 Mouse spl
	31	63.2	5.0	549	10	ADK58517	Adk58517 Plant DNA
	32	57.8	4.6	1196	12	ADJ39614	Adj39614 Plant cDN
	33	53.8	4.3	1364	12	ADJ42189	Adj42189 Plant cDN
	34	47.4	3.8	1747	6	AAS94774	Aas94774 Human DNA
	35	47.4	3.8	1863	3	AAZ47544	Aaz47544 Human pro
	36	47.4	3.8	2044	8	ABX70932	Abx70932 Novel hum
	37	47.4	3.8	2887	12	ADJ96566	Adj96566 Human DYR
	38	45.4	3.6	1581	6	ABK88905	Abk88905 Murine mD
	39	45.2	3.6	7565	12	ADK70703	Adk70703 Murine CR
	40	44.4	3.5	2658	3	AAA70213	Aaa70213 Plasmodiu
	41	43.8	3.5	1584	6	ABL58336	Abl58336 Rat YAK1
	42	43	3.4	681	10	ADK52877	Adk52877 Plant DNA
c	43	42.8	3.4	600	12	ADJ44008	Adj44008 Plant cDN
	44	42.4	3.4	1497	3	AAA70154	Aaa70154 Plasmodiu
	45	42.2	3.3	2095	3	AAC36941	Aac36941 Arabidops

ALIGNMENTS

RESULT 1
AAZ51355
ID AAZ51355 standard; DNA; 1260 BP.
XX
AC AAZ51355;
XX
DT 06-JUN-2000 (first entry)
XX
DE Human wild type serine/threonine kinase KIS (hKIS) gene.
XX
KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
KW dominant gene; cyclin-dependent kinase inhibitor; CKI; p27; modulator;
KW treatment; cell proliferative disease; vascular disorder; gene therapy;
KW atherosclerosis; restenosis; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1260
FT /*tag= a
FT /product= "Human wild type KIS protein"
FT /transl_except= (pos:556. .561, Asn)
XX
PN WO200011165-A1.
XX
PD 02-MAR-2000.
XX
PF 20-AUG-1999; 99WO-US018903.
XX
PR 21-AUG-1998; 98US-0097710P.
XX
PA (NABE/) NABEL G J.
PA (NABE/) NABEL E G.
XX
PI Nabel GJ, Nabel EG;
XX
DR WPI; 2000-237648/20.
DR P-PSDB; AAY70305.
XX
PT Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT proliferation.
XX
PS Claim 1; Page 56; 70pp; English.

```
XX
CC   The present DNA sequence encodes a wild type human KIS (hKIS), isolated
CC   from a human B-cell library, using a yeast two hybrid screening system.
CC   hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
CC   cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
CC   that controls cell proliferation and is localised predominantly in the
CC   nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC   G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
CC   The hKIS sequences are used to modulate cell proliferation and treat cell
CC   proliferative and vascular diseases. The polynucleotide sequence may be
CC   used in gene therapy to treat vascular disorders such as restenosis or
CC   atherosclerosis
XX
SQ   Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 1260;  DB 3;  Length 1260;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1260;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260


RESULT 2
AAZ51356
ID   AAZ51356 standard; DNA; 1260 BP.
XX
AC   AAZ51356;
XX
DT   06-JUN-2000  (first entry)
XX
DE   Human transdominant mutant serine/threonine kinase KIS (hKIS) gene.
XX
KW   KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
KW   transdominant gene; mutant; cyclin-dependent kinase inhibitor; CKI; p27;
KW   modulator; treatment; cell proliferative disease; vascular disorder;
KW   gene therapy; atherosclerosis; restenosis; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1260
FT                   /*tag= a
FT                   /product= "Human transdominant mutant KIS protein"
FT                   /transl_except= (pos:556. .561, Asn)
FT   misc_feature    161
FT                   /*tag= b
FT                   /note= "A to G substitution resulting in a Lys to Arg
FT                   change in the amino acid sequence"
XX
PN   WO200011165-A1.
XX
PD   02-MAR-2000.
XX
PF   20-AUG-1999;   99WO-US018903.
XX
PR   21-AUG-1998;   98US-0097710P.
XX
PA   (NABE/) NABEL G J.
PA   (NABE/) NABEL E G.
XX
PI   Nabel GJ,  Nabel EG;
XX
DR   WPI; 2000-237648/20.
DR   P-PSDB; AAY70306.
```

XX
PT Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT proliferation.
XX
PS Claim 7; Page 59; 70pp; English.
XX
CC The present DNA sequence encodes a transdominant mutant human KIS (hKIS),
CC constructed by site directed mutagenesis. A single nucleotide
CC substitution (A to G) results in a lysine to arginine change in the
CC protein sequence. hKIS is a serine/threonine kinase, that acts as an
CC inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
CC controls cell proliferation and is localised predominantly in the
CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC G1 phase. The hKIS sequences are used to modulate cell proliferation and
CC treat cell proliferative and vascular diseases. The polynucleotide
CC sequence may be used in gene therapy to treat vascular disorders such as
CC restenosis or atherosclerosis
XX
SQ Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;

Query Match 99.9%; Score 1258.4; DB 3; Length 1260;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

RESULT 3
ABQ99383
ID ABQ99383 standard; cDNA; 2008 BP.
XX
AC ABQ99383;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human coding sequence SEQ ID 116.
XX
KW Human; expressed sequence tag; EST; haematopoietic disorder;
KW central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200259260-A2.
XX
PD 01-AUG-2002.
XX
PF 16-NOV-2001; 2001WO-US042950.
XX
PR 17-NOV-2000; 2000US-00714936.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-590824/63.
DR N-PSDB; ABP64797.

```
XX
PT   New isolated polynucleotide, useful in research, diagnostic or
PT   therapeutic methods, e.g. preventing or treating disorders involving
PT   aberrant protein expression or biological activity.
XX
PS   Claim 1; SEQ ID NO 116; 394pp; English.
XX
CC   The present invention relates to novel human coding sequences (ABQ99268-
CC   ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC   therapeutic, diagnostic and research methods. The polynucleotides may be
CC   used in the field of molecular biology as hybridisation probes, primers
CC   for PCR, for chromosome and gene mapping, for the recombinant production
CC   of protein, or in generation of anti-sense DNA or RNA. The
CC   polynucleotides are useful in diagnostics as expressed sequence tags
CC   (ESTs) for identifying expressed genes or for physical mapping of the
CC   human genome. The proteins may be used as molecular weight markers, or as
CC   nutritional sources or supplements. The proteins may be used to maintain
CC   and expand cell population in a totipotential or pluripotential state
CC   useful for re-engineering damaged or diseased tissues, transplantation,
CC   manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC   polynucleotides and proteins are useful for preventing, treating or
CC   ameliorating disorders involving aberrant protein expression or
CC   biological activity, e.g. haematopoietic disorders, central/peripheral
CC   nervous system diseases, mechanical and traumatic disorders, non-healing
CC   wounds, immune deficiencies and disorders, infectious diseases caused by
CC   viral, bacterial or fungal infection, autoimmune disorders, allergic
CC   reactions and conditions, coagulation disorders, or cancer. The
CC   polynucleotide sequences of the invention were assembled from ESTs
CC   isolated mainly by sequencing by hybridisation, and in some cases,
CC   sequences obtained from one or more public databases. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;

  Query Match             99.9%;  Score 1258.4;  DB 6;  Length 2008;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1259;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 138

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 198

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 258

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 318

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 378

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 438

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 498

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 558

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 618

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 678

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 738

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 798

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 858

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 918

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 978

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        979 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1038

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1039 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1098

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1099 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1158

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1159 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1218

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1219 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1278

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1279 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1338

RESULT 4
AAI66829
ID   AAI66829 standard; cDNA; 1260 BP.
XX
AC   AAI66829;
XX
DT   07-JAN-2002  (first entry)
XX
DE   Human protein kinase polypeptide 13887 coding sequence.
XX
KW   Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KW   cytostatic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1260
FT                   /*tag=  a
FT                   /product= "protein kinase 13887"
XX
```

PN WO200173050-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-US009483.
XX
PR 24-MAR-2000; 2000US-0191846P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers R;
XX
DR WPI; 2001-611632/70.
DR P-PSDB; AAG65767.
XX
PT New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT useful in diagnosis of cancer or cellular proliferation or
PT differentiation disorders and to screen for polypeptide modulators useful
PT to treat such conditions.
XX
PS Claim 1; Fig 10A-C; 169pp; English.
XX
CC The invention provides novel human protein kinase polypeptides, 3714,
CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC protein kinase polypeptides can be expressed by standard recombinant
CC methodology.3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC are useful for diagnostic and screening methods to identify subjects (at
CC risk of) having cancer or cellular proliferation and/or differentiation
CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC modulators are useful for the treatment of cancer, particularly colon
CC cancer or cellular proliferation and/or differentiation disorders. Other
CC disorders associated with 3714, 16742, 23546 or 13887 expression or
CC activity that can be treated include bone related disorders, inflammatory
CC disorders, autoimmune diseases, cardiovascular disorders and liver
CC diseases. The present sequence represents a human protein kinase
CC polypeptide 13887 coding sequence
XX
SQ Sequence 1260 BP; 312 A; 279 C; 334 G; 335 T; 0 U; 0 Other;

Query Match 99.7%; Score 1256.8; DB 4; Length 1260;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

RESULT 5
AAI66828
ID AAI66828 standard; cDNA; 2622 BP.
XX
AC AAI66828;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human protein kinase polypeptide 13887 encoding cDNA.
XX
KW Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KW cytostatic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1. .266

```
FT                   /*tag=  a
FT   CDS             267. .1529
FT                   /*tag=  b
FT                   /product= "protein kinase 13887"
FT                   /note= "coding sequence specifically claimed"
FT   3'UTR           1530. .2622
FT                   /*tag=  c
XX
PN   WO200173050-A2.
XX
PD   04-OCT-2001.
XX
PF   23-MAR-2001; 2001WO-US009483.
XX
PR   24-MAR-2000; 2000US-0191846P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Meyers R;
XX
DR   WPI; 2001-611632/70.
DR   P-PSDB; AAG65767.
XX
PT   New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT   useful in diagnosis of cancer or cellular proliferation or
PT   differentiation disorders and to screen for polypeptide modulators useful
PT   to treat such conditions.
XX
PS   Claim 1; Fig 10A-C; 169pp; English.
XX
CC   The invention provides novel human protein kinase polypeptides, 3714,
CC   16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC   protein kinase polypeptides can be expressed by standard recombinant
CC   methodology.3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC   are useful for diagnostic and screening methods to identify subjects (at
CC   risk of) having cancer or cellular proliferation and/or differentiation
CC   disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC   modulators are useful for the treatment of cancer, particularly colon
CC   cancer or cellular proliferation and/or differentiation disorders. Other
CC   disorders associated with 3714, 16742, 23546 or 13887 expression or
CC   activity that can be treated include bone related disorders, inflammatory
CC   disorders, autoimmune diseases, cardiovascular disorders and liver
CC   diseases. The present sequence represents a human protein kinase
CC   polypeptide 13887 encoding cDNA
XX
SQ   Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other;

  Query Match           99.7%;  Score 1256.8;  DB 4;  Length 2622;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1258;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 326

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 386

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 446

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 506

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 566

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        567 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 626

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 686

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 746

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 806

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 866

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 926

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 986

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1046

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1106

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1166

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1226

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1286

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1346

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1406

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1466

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1526


RESULT 6
AAH72714
ID   AAH72714 standard; cDNA; 4065 BP.
XX
AC   AAH72714;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 3988.
XX
```

```
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer and
PT   for assessing and detecting compounds for treating the cancer.
XX
PS   Claim 1; Page 810-811; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

  Query Match           99.7%;  Score 1256.8;  DB 4;  Length 4065;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1258;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1527


RESULT 7
ADL45238
ID   ADL45238 standard; DNA; 4065 BP.
XX
AC   ADL45238;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human ovarian cancer DNA marker #19128.
XX
KW   Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
OS   Homo sapiens.
XX
PN   WO200170979-A2.
XX
PD   27-SEP-2001.
XX
```

PF 21-MAR-2001; 2001WO-US009126.
XX
PR 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lee J, Lillie J;
XX
DR WPI; 2001-611502/70.
XX
PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX
PS Disclosure; SEQ ID NO 19128; 106pp; English.
XX
CC The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.
XX
SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

Query Match 99.7%; Score 1256.8; DB 5; Length 4065;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1527
```

RESULT 8
AAS06734

```
ID   AAS06734 standard; cDNA; 1260 BP.
XX
AC   AAS06734;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Polynucleotide sequence encoding human protein kinase #34.
XX
KW   Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW   metabolic disorder; immune related disease; neurological disorder;
KW   neurodegenerative disorder; inflammatory disorder; infectious disease;
KW   reproductive disorder; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200138503-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US032085.
XX
PR   24-NOV-1999;   99US-0167482P.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman GD,  Whyte D,  Manning G,  Sudarsanam S,  Martinez R;
PI   Flanagan P,  Clary D;
XX
DR   WPI; 2001-343950/36.
DR   P-PSDB; AAU03534.
XX
PT   Nucleic acids encoding human kinase polypeptides, useful for preventing
PT   diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT   neuronal-associated diseases, and microbial infections.
XX
PS   Example 1; Fig 1; 433pp; English.
XX
CC   AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
CC   protein kinases have been identified as members of the tyrosine or
CC   serine/threonine kinase (PTK and STK) families. The polynucleotides
CC   encoding protein kinases and the polypeptides may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate kinase expression. For example, they may be used to treat
CC   cancers (especially cancers of haematopoietic origin), cardiovascular
CC   disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC   immune related diseases (e.g. rheumatoid arthritis), neurological
CC   disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC   Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC   disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC   Additionally, polynucleotides encoding protein kinases may be used for
CC   gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC   polypeptides may be used as antigens in the production of antibodies
CC   against the protein kinases and in assays to identify modulators of
CC   protein kinase expression and activity
XX
SQ   Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other;

  Query Match             99.6%;  Score 1255.2;  DB 4;  Length 1260;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1257;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

RESULT 9
AAD18818
ID AAD18818 standard; cDNA; 1824 BP.
XX
AC AAD18818;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human kinase (PKIN)-3 cDNA.
XX
KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;
KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
KW antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 169. .1428
FT /*tag= a
FT /product= "Human PKIN-3 protein"
XX
PN WO200181555-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012992.
XX
PR 20-APR-2000; 2000US-0199021P.
PR 28-APR-2000; 2000US-0200226P.
PR 05-MAY-2000; 2000US-0202339P.
PR 11-MAY-2000; 2000US-0203505P.
PR 18-MAY-2000; 2000US-0205564P.
PR 26-MAY-2000; 2000US-0207739P.
PR 01-JUN-2000; 2000US-0208795P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
PI Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Walia NK, Yao MG;
PI Patterson C, Burrill JD, Marcus GA, Zingler KA, Recipon SA, Lu Y;
PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
PI Walsh RT, Ramkumar J, Borowsky ML, Au-Young J, Hillman JL;
PI Gururajan R;
XX
DR WPI; 2001-611740/70.
DR P-PSDB; AAE11769.
XX
PT Human kinases and nucleic acids, useful for preventing diagnosing and
PT treating cancers, inflammation and immune disorders.
XX
PS Claim 5; Page 152; 166pp; English.
XX
CC The present invention relates to human kinases (PKIN) and the nucleic
CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
CC used in the prevention, diagnosis and treatment of diseases cancers,
CC adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
CC gout, microbial infections, cardiovascular disease and/or inflammation,
CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
CC infarction, cataract, growth and development disorder, seizure disorder,
CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
CC PKIN may be used to treat disorders associated with decreased PKIN
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of PKIN by expressing inactive proteins or to
CC supplement the patients own production of PKIN. PKIN nucleic acids may be
CC used to produce the PKIN polypeptide, by inserting the nucleic acids into
CC a host cell and culturing the cell to express the protein. PKIN nucleic
CC acid and its complementary sequences may also be used as DNA probes in
CC diagnostic assays to detect and quantitate the presence of similar
CC nucleic acid sequences in samples and therefore which patients may be in
CC need of restorative therapy. The present sequence is human PKIN-3 cDNA
XX
SQ Sequence 1824 BP; 434 A; 416 C; 472 G; 502 T; 0 U; 0 Other;

Query Match 99.6%; Score 1255.2; DB 4; Length 1824;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 228

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 288

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 348

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 408

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 468

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 528

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        529 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 588

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 648

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 708

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 768

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 828

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 888

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 948

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1008

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1068

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1128

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
                ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1129 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1188

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1189 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1248

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
                ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1249 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1308

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1309 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1368

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1369 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1428
```

RESULT 10
ADQ24401
ID ADQ24401 standard; DNA; 3998 BP.
XX
AC ADQ24401;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7221.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 7221; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 3998 BP; 1066 A; 777 C; 885 G; 1255 T; 0 U; 15 Other;

  Query Match   93.3%; Score 1175; DB 12; Length 3998;
  Best Local Similarity 96.2%; Pred. No. 0;
  Matches 1215; Conservative   0; Mismatches   45; Indels   3; Gaps   1;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 219

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 279

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 339

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||||||||||||||||||||||||  |||          |  | |
Db        340 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTAAGCATGTTTCTTACCAGACCG 399

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              |  |  |  | |  |  | |||||||||||||||||||||||||||||  ||
Db        400 AAGGTCTGTGTTGACCTTAACCGGAGAGTGACTTTGTATGGAGTGTTTACAATCCAGGTT 459

Qy        301 TCT---CCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCG 357
                |   | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 GTTCTCCAAAGTGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCG 519

Qy        358 GAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCC 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        520 GAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCC 579

Qy        418 CGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTC 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 CGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTC 639

Qy        478 AAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGA 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 AAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGA 699

Qy        538 CTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCT 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 CTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCT 759

Qy        598 CCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGT 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 CCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGT 819

Qy        658 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 879

Qy        718 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 939

Qy        778 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 999

Qy        838 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 1059

Qy        898 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCC 1119

Qy        958 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAA 1017
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
```

```
Db       1120 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAA 1179

Qy       1018 TATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCT 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 TATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCT 1239

Qy       1078 CTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCT 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 CTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCT 1299

Qy       1138 GGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 GGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTT 1359

Qy       1198 GTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1360 GTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1419

Qy       1258 TAA 1260
              |||
Db       1420 TAA 1422


RESULT 11
ADB52857
ID   ADB52857 standard; DNA; 3244 BP.
XX
AC   ADB52857;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3399.
XX
KW   toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW   toxicity marker; toxicity progression; drug screening;
KW   primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS   Rattus norvegicus.
XX
PN   WO2003065993-A2.
XX
PD   14-AUG-2003.
XX
PF   04-FEB-2003; 2003WO-US003482.
XX
PR   04-FEB-2002; 2002US-0353171P.
PR   13-MAR-2002; 2002US-0363534P.
PR   08-APR-2002; 2002US-0370248P.
PR   10-APR-2002; 2002US-0371134P.
PR   10-APR-2002; 2002US-0371135P.
PR   10-APR-2002; 2002US-0371150P.
PR   11-APR-2002; 2002US-0371413P.
PR   19-APR-2002; 2002US-0373601P.
PR   19-APR-2002; 2002US-0373602P.
PR   22-APR-2002; 2002US-0374139P.
PR   08-MAY-2002; 2002US-0378370P.
PR   09-MAY-2002; 2002US-0378652P.
PR   09-MAY-2002; 2002US-0378653P.
PR   09-MAY-2002; 2002US-0378665P.
PR   09-JUL-2002; 2002US-0394230P.
PR   09-JUL-2002; 2002US-0394253P.
PR   04-SEP-2002; 2002US-0407688P.
PR   28-JAN-2003; 2003US-0442900P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,  Porter M,  Johnson K,  Higgs B,  Castle A,  Orr M;
PI   Elashoff M;
XX
DR   WPI; 2003-731472/69.
XX
PT   Determining if a compound induces a toxic effect on a tissue or cell, for
PT   identifying hepatotoxic compounds, comprises comparing a gene expression
PT   profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT   mean values.
XX
PS   Claim 44; SEQ ID NO 3399; 874pp; English.
XX
CC   The present invention describes a method for determining whether a
CC   compound induces a toxic effect on a tissue or cell. The method comprises
CC   preparing a gene expression profile of a tissue or cell sample exposed to
CC   the compound, and comparing the gene expression profile to a database
CC   comprising data or information on the Tox mean and non-Tox mean value.
CC   The method is useful for predicting or identifying at least one toxic
CC   effect, particularly hepatotoxicity, of a test or unknown compound. The
CC   genes listed in the specification are useful as diagnostic or toxicity
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 3244 BP; 800 A; 710 C; 819 G; 915 T; 0 U; 0 Other;

  Query Match             87.9%;  Score 1108;  DB 10;  Length 3244;
  Best Local Similarity   92.5%;  Pred. No. 0;
  Matches 1165;  Conservative    0;  Mismatches   95;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||| || |||||||||| |||||||| |||||||||| || ||||||||||||||||
Db        162 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 221

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              |||||||||||||||||||| || || |||||||||| |||||||||| |||||
Db        222 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 281

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||  || |||||||||||||||||||||||||||| |||| || |||
Db        282 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 341

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||| ||||| || || |||||||||||||| |||||||||| ||||||
Db        342 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 401

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              || |||||||||||||||| |||||||||||||| || ||||| ||||| || |||||
Db        402 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 461

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        462 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 521

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||  |||||||||||||||||||| |||||||||||||| ||||| || ||| ||
Db        522 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 581

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||  |||||||||||||||||||| |||||||||||||||||| ||||||||
Db        582 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 641

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||| |||||  |||||||||| ||||| || || ||||| || ||||||||||||||
Db        642 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 701

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||| || ||||||||||||||||||||||||||||| |||||
Db        702 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 761

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||| |||| |||||||||||||||| |||||||||||||||||||| ||||||
Db        762 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 821

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||| |||||||| |||||||||||||||||||||||||||||||||||
Db        822 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 881
```

```
Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        882 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 941

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        942 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 1001

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              |||||||||||||||||||| || |||||||||| || |||||||| ||||| ||||||
Db       1002 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 1061

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||| |||
Db       1062 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 1121

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||| ||||| || |||||||||||||||||| ||||||| |||||||| ||||||
Db       1122 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1181

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              |||||||||||||||||||||||| |||||||| |||||||||||||||||| |||||
Db       1182 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1241

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db       1242 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1301

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              |||||||||||||| |||||||| |||||||| |||||||||| |||||||||||||||||
Db       1302 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1361

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1362 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1421
```

RESULT 12
AAH70764
ID AAH70764 standard; cDNA; 529 BP.
XX
AC AAH70764;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 2038.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 434; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

```
  Query Match             36.4%;  Score 458.6;  DB 4;  Length 529;
  Best Local Similarity   97.8%;  Pred. No. 3.5e-125;
  Matches  486;  Conservative    0;  Mismatches    9;  Indels    2;  Gaps    2;

Qy        654 ATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 713
              | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 AGGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 81

Qy        714 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 773
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 141

Qy        774 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 833
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 201

Qy        834 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 893
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 261

Qy        894 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 953
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 321

Qy        954 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGA 1013
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        322 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGA 381

Qy       1014 GGAATATGAAGA-TGTTGTAGAAGATGTAAAAGA-GGAGTGTCAAAAATATGGACCAGTG 1071
              |||||||||||| |||||||||||||||||||| |||||||||||||||||||||||| |
Db        382 GGAATATGAAGATTGTTGTAGAAGATGTAAAAGAGGGAGTGTCAAAAATATGGACCAGGG 441

Qy       1072 GTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCA 1131
              |||||||||||||| |||||||||||||||||||||||||  ||||||||||||||||||
Db        442 GTATCTCTACTTGGTCCAAAGGAAAATCCTGGCAGAGGAACAGTCTTTGTTGAGTATGCA 501

Qy       1132 AATGCTGGTGATTCCAA 1148
              || ||||| ||||| ||
Db        502 AAGGCTGGGGATTCAAA 518
```

RESULT 13
AAA02535
ID AAA02535 standard; cDNA; 722 BP.
XX
AC AAA02535;
XX
DT 19-MAY-2000 (first entry)
XX
DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:2526.
XX
KW Human; colon cancer; tumour; diagnosis; gene expression product; probe;
KW detection; cancerous state; metastasis; identification; breast cancer;
KW oestrogen receptor-positive breast cancer; therapy;
KW oestrogen receptor-negative breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
PN WO9958675-A2.

```
XX
PD   18-NOV-1999.
XX
PF   13-MAY-1999;   99WO-US010602.
XX
PR   14-MAY-1998;   98US-0085426P.
PR   15-MAY-1998;   98US-0085537P.
PR   15-MAY-1998;   98US-0085696P.
PR   21-OCT-1998;   98US-0105234P.
PR   27-OCT-1998;   98US-0105877P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Williams LT,  Escobedo J,  Innis MA,  Garcia PD,  Sudduth-Klinger J;
PI   Reinhard C,  Giese K,  Randazzo F,  Kennedy GC,  Pot D,  Kassam A;
PI   Lamson G,  Drmanac R,  Crkvenjakov R,  Dickson M,  Drmanac S,  Labat I;
PI   Leshkowitz D,  Kita D,  Garcia V,  Jones LW,  Stache-Crain B;
XX
DR   WPI; 2000-126369/11.
XX
PT   Polynucleotide library used to determine cancerous states of mammalian
PT   cells.
XX
PS   Claim 1; Page 1020; 1097pp; English.
XX
CC   AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC   libraries constructed from human colon cancer cell lines. The present
CC   invention also describes a method of detecting differentially expressed
CC   genes correlated with a cancerous state of a mammalian cell, comprising
CC   detecting at least one differentially expressed gene product in a test
CC   sample derived from a cell suspected of being cancerous, where detection
CC   of the differentially expressed gene product is correlated with a
CC   cancerous state of the cell from which the test sample was derived. The
CC   polynucleotides sequences can be used in a method for detecting
CC   differentially expressed genes correlated with a cancerous state of a
CC   mammalian cell. The polynucleotides can also be used as probes for
CC   detecting and mapping related genes. They can be used in diagnosis and
CC   prognosis of diseases and disorders (e.g. identification of pre-
CC   metastatic or metastatic cancerous states, stages of cancer, or
CC   responsiveness of cancer to therapy). This is particularly for breast
CC   cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC   negative breast cancer, lung cancer, and colon cancer
XX
SQ   Sequence 722 BP; 215 A; 129 C; 164 G; 193 T; 0 U; 21 Other;

  Query Match             35.4%;  Score 445.8;  DB 3;  Length 722;
  Best Local Similarity   93.0%;  Pred. No. 2.7e-121;
  Matches  581;  Conservative    0;  Mismatches   29;  Indels   15;  Gaps   11;

Qy        454 GAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAA 513
              |||||||||||||||||| |||||||| |||||||||||||||||||||||||||||||
Db          1 GAGGGCTATGTCCATGCGGNCCTCAAACNACGTAACATATTGTGGAGTGCAGAGAATGAA 60

Qy        514 TGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 573
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db         61 TGTTTTAAACTCATTGACTTTGGACTTANCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 120

Qy        574 ATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCT 633
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||| |||
Db        121 ATTCAGACAGACGGGTATCGGGCTCCAGAA-CAGAATTGCAAAATTGCTTGGCCCANGCT 179

Qy        634 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 693
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 239

Qy        694 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 753
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 299

Qy        754 GCAAACAG-TTCTGCTATTA-TTGATCACATATTTGCCAGTAAA--GCAGTGGTGAATGC 809
              |||||||| |||||||||| |||||||||||||||||||||    | ||||||||||
Db        300 GCAAACAGTTTCTGCTATTATTTGATCACATATTTGCCAGTAAAAGCAANTGGTGAATGC 359

Qy        810 CGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCC--A 867
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||  |
Db        360 CGCAATTCCAGCCTATCACCTAANAGACCTTATCAAAAGCATGCTTCATGATGATCCCAA 419

Qy        868 AGCAGAAGAATTCCTGCTGAAATGGCATT-GTGCAGCCCATTCTTTAGC-ATTCCTTTTG 925
              |||||||| || |||||||||||||| ||||| |||||||||||||| |||||||||
Db        420 GCAGGAAGAATTTCTNCTGAAATGGCATTGGTGCANCCCATTCTTTAGCNATTCCTTTTG 479

Qy        926 CCCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGT-GCTAAGACTGCTGAATGTG 984
              |||||||||||||||||| |||||||| ||||||||| ||||||||||||||||||||
Db        480 CCCCTCATATTGAAGATCTGNTCATGCTTTCCACTCCAGTGGCTAAGACTGCTGAATGTG 539

Qy        985 CT-GGATGATGATTATCTTGGGAATGAA--GAGGAATATGAAGATGTTGTAGAAGATGT- 1040
              || || |||||||||||||| |||||||   |||| |||||||||||||| |||||||
Db        540 CTGGGNTGATGATTATCTTGAGAATGAAAGAAGGATTATGAAGATGTTGTTGAAGATGNT 599

Qy       1041 -AAAAGAGGAGTGTCAAAAATATGG 1064
               |||  || |||| ||||||| | |
Db        600 AAAAGAAGAAGTGGCAAAAATNTTG 624


RESULT 14
AAH71966
ID   AAH71966 standard; cDNA; 470 BP.
XX
AC   AAH71966;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 3240.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer and
PT   for assessing and detecting compounds for treating the cancer.
XX
PS   Claim 1; Page 636; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;

  Query Match             33.9%;  Score 426.6;  DB 4;  Length 470;
  Best Local Similarity   98.7%;  Pred. No. 1e-115;
```

```
  Matches  451;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy        654 ATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 713
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 AGGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 73

Qy        714 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 133

Qy        774 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 193

Qy        834 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 253

Qy        894 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 313

Qy        954 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGA 1013
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        314 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGA 373

Qy       1014 GGAATATGAAGA-TGTTGTAGAAGATGTAAAAGA-GGAGTGTCAAAAATATGGACCAGTG 1071
              |||||||||||| |||||||||||||||||||| |||||||||||||||||||||||| |
Db        374 GGAATATGAAGATTGTTGTAGAAGATGTAAAAGAGGGAGTGTCAAAAATATGGACCAGGG 433

Qy       1072 GTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAG 1108
              |||||||||||| |||||||||||||||||||||||
Db        434 GTATCTCTACTTGGTCCAAAGGAAAATCCTGGCAGAG 470


RESULT 15
AAH69101
ID   AAH69101 standard; cDNA; 461 BP.
XX
AC   AAH69101;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 375.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer and
PT   for assessing and detecting compounds for treating the cancer.
XX
PS   Claim 1; Page 168-169; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 461 BP; 134 A; 91 C; 109 G; 127 T; 0 U; 0 Other;

  Query Match             29.8%;  Score 375.4;  DB 4;  Length 461;
  Best Local Similarity   96.2%;  Pred. No. 1.6e-100;
  Matches  428;  Conservative    0;  Mismatches   11;  Indels    6;  Gaps    4;

Qy        654 ATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 713
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 AGGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 77

Qy        714 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 137

Qy        774 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 197

Qy        834 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 257

Qy        894 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 953
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        258 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATA-TGAAGATCTGGTCATGCT 316

Qy        954 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGA 1013
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        317 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGA 376

Qy       1014 GGAATATGAAGAT--GTTGTAGAAGATGTAAAAGAGG--AGTGTCAAAAATATGGACCA- 1068
              |||||||||||||    |||||||||||||||||||   |||||||||||||||||||| |
Db        377 GGAATATGAAGATTGTTGTTAGAAGATGTAAAAGAGGGAGGTGTCAAAAATATGGACCAG 436

Qy       1069 GTGGTATCTCTACTTGTTCCAAAGG 1093
              ||||||||||||||  | | || |
Db        437 GTGGTATCTCTACTTTGTTCCAAAG 461


Search completed: November 29, 2004, 11:45:24
Job time : 656.5 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 4563 Seconds
(without alignments)
4342.403 Million cell updates/sec

Title: US-10-798-532-2
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAFG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10798532/runat_29112004_112452_15890/app_query.fasta_1.1166
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532_@CGN_1_1_6317_@runat_29112004_112452_15890 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2198	99.6	1260	6	CQ715025	CQ715025 Sequence
2	2198	99.6	1260	6	AX262521	AX262521 Sequence
3	2198	99.6	2598	6	AX262519	AX262519 Sequence
4	2198	99.6	2909	9	HSA536197	AJ536197 Homo sapi
5	2198	99.6	4065	6	CQ412057	CQ412057 Sequence
6	2198	99.6	4065	6	AX188296	AX188296 Sequence
7	2191	99.3	1260	6	AX166543	AX166543 Sequence
8	2191	99.3	1824	6	AX680136	AX680136 Sequence
9	2186	99.1	1633	10	RNKISRNA	X98374 R.norvegicu
10	2186	99.1	3244	10	RNU70372	U70372 Rattus norv
11	2185	99.0	2041	10	BC058732	BC058732 Mus muscu
c 12	2177	98.7	110000	2	AC129141_1	Continuation (2 of
13	2160	97.9	1254	10	MMPKKIS	Y10725 Mus musculu
14	1993	90.3	1367	9	BC026046	BC026046 Homo sapi
15	1893.5	85.8	2940	9	HSM805327	AL834136 Homo sapi
16	1764.5	80.0	1157	5	CR386837	CR386837 Gallus ga
17	1721	78.0	1135	9	AK058195	AK058195 Homo sapi
18	1634	74.1	2041	9	BC014917	BC014917 Homo sapi
c 19	1050	47.6	259562	2	AC094499	AC094499 Rattus no
20	1045	47.4	601	10	MMKISRNA	X82320 M.musculus
21	908	41.2	105902	9	AL359699	AL359699 Human DNA
22	908	41.2	167995	2	AC040901	AC040901 Homo sapi
c 23	904	41.0	211812	2	AC027205	AC027205 Homo sapi
24	805.5	36.5	211635	2	AC112763	AC112763 Rattus no
25	789	35.8	529	6	AX186343	AX186343 Sequence
26	766.5	34.7	722	6	BD221413	BD221413 Human gen
27	729	33.0	470	6	AX187546	AX187546 Sequence
28	724	32.8	100207	10	AY180177	AY180177 Mus muscu
c 29	697.5	31.6	200095	10	AC121606	AC121606 Mus muscu
30	697.5	31.6	220806	10	AC123650	AC123650 Mus muscu
31	663	30.1	462	6	AX184680	AX184680 Sequence
32	369	16.7	60640	10	AY180176	AY180176 Mus muscu
33	369	16.7	185810	2	AC136052	AC136052 Rattus no
34	361	16.4	231	6	BD220137	BD220137 Human gen
35	330.5	15.0	205752	2	AC119431	AC119431 Mus muscu
36	311.5	14.1	419	6	CQ393016	CQ393016 Sequence
37	311.5	14.1	419	6	CQ399411	CQ399411 Sequence
c 38	267	12.1	185810	2	AC136052	AC136052 Rattus no
39	257	11.7	110000	8	CR382128_04	Continuation (5 of
40	252	11.4	1065	10	HAMCDK2A	D17350 Mesocricetu
41	250.5	11.4	1269	10	AB096078	AB096078 Mus muscu
42	250.5	11.4	1395	10	AB096079	AB096079 Mus muscu
43	249.5	11.3	1400	10	RATCDK2A	D28753 Rattus ratt
44	249.5	11.3	2250	10	BC061832	BC061832 Rattus no
45	248.5	11.3	2809	3	AY064464	AY064464 Caenorhab

ALIGNMENTS

RESULT 1
CQ715025
LOCUS CQ715025 1260 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 959 from Patent WO02068579.
ACCESSION CQ715025
VERSION CQ715025.1 GI:42275882
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 959 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1. .1260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.: 2.81e-191 Length: 1260
Score: 2198.00 Matches: 418

```
Percent Similarity:     99.76%          Conservative:  0
Best Local Similarity:  99.76%          Mismatches:    1
Query Match:            99.64%          Indels:        0
DB:                     6               Gaps:          0

US-10-798-532-2 (1-419) x CQ715025 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
```

```
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257
```

RESULT 2
AX262521

```
LOCUS       AX262521                1260 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 12 from Patent WO0173050.
ACCESSION   AX262521
VERSION     AX262521.1  GI:16511408
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Meyers,R.
  TITLE     3714, 16742, 23546, and 13887 novel protein kinase molecules and
            uses therefor
  JOURNAL   Patent: WO 0173050-A 12 04-OCT-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1260
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.81e-191       Length:         1260
Score:                  2198.00         Matches:        418
Percent Similarity:     99.76%          Conservative:   0
Best Local Similarity:  99.76%          Mismatches:     1
Query Match:            99.64%          Indels:         0
DB:                     6               Gaps:           0

US-10-798-532-2 (1-419) x AX262521 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
```

```
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140

Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160

Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180

Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200

Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220

Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240

Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260

Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280

Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300

Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320

Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340

Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360

Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400

Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419

Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257
```

RESULT 3
AX262519

```
LOCUS       AX262519               2598 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 10 from Patent WO0173050.
ACCESSION   AX262519
VERSION     AX262519.1  GI:16511406
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Meyers,R.
  TITLE     3714, 16742, 23546, and 13887 novel protein kinase molecules and
            uses therefor
  JOURNAL   Patent: WO 0173050-A 10 04-OCT-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2598
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             268. .1527
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD10238.1"
                     /db_xref="GI:16511407"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPPAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN
```

Alignment Scores:

Pred. No.:	7.7e-191	Length:	2598
Score:	2198.00	Matches:	418
Percent Similarity:	99.76%	Conservative:	0
Best Local Similarity:	99.76%	Mismatches:	1
Query Match:	99.64%	Indels:	0
DB:	6	Gaps:	0

US-10-798-532-2 (1-419) x AX262519 (1-2598)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20

Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40

Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60

Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80

Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100

Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120

Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140

Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160

Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180

Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807
```

```
Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1524
```

RESULT 4
HSA536197

```
LOCUS       HSA536197               2909 bp    mRNA    linear   PRI 11-JAN-2003
DEFINITION  Homo sapiens mRNA for KIS protein.
ACCESSION   AJ536197
VERSION     AJ536197.1  GI:27657360
KEYWORDS    KIS protein.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Bieche,I., Manceau,V., Curmi,P.A., Laurendon,I., Lachkar,S.,
            Leroy,K., Vidaud,D., Sobel,A. and Maucuer,A.
  TITLE     Quantitative RT-PCR reveals a ubiquitous but preferentially neural
            expression of the KIS gene in rat and human
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2909)
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-2003) Maucuer A., U440, Inserm, IFM, 17 rue du
            fer a moulin, 75005 Paris, FRANCE
FEATURES             Location/Qualifiers
     source          1. .2909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:IMAGp998M213680"
                     /db_xref="taxon:9606"
                     /clone="I.M.A.G.E. Consortium CloneID 1455548"
                     /tissue_type="2 pooled tumors (clear cell type)"
     gene            1. .2909
                     /gene="KIS"
     CDS             146. .1405
                     /gene="KIS"
                     /codon_start=1
                     /product="KIS protein"
                     /protein_id="CAD60192.1"
                     /db_xref="GI:27657361"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALALFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN
```

Alignment Scores:

Pred. No.:	9.01e-191	Length:	2909
Score:	2198.00	Matches:	418
Percent Similarity:	99.76%	Conservative:	0
Best Local Similarity:	99.76%	Mismatches:	1
Query Match:	99.64%	Indels:	0
DB:	9	Gaps:	0

US-10-798-532-2 (1-419) x HSA536197 (1-2909)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 205

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 265

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 325

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 385

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 445

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 505

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 565

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 625

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CCACGTAACATATTGTGGAGTCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 685

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 745
```

```
Qy       201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
Db       746 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 805

Qy       221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
Db       806 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 865

Qy       241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
Db       866 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 925

Qy       261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
Db       926 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 985

Qy       281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
Db       986 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1045

Qy       301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
Db      1046 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1105

Qy       321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
Db      1106 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1165

Qy       341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
Db      1166 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1225

Qy       361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
Db      1226 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1285

Qy       381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Db      1286 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1345

Qy       401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db      1346 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1402
```

```
RESULT 5
CQ412057
LOCUS       CQ412057                4065 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 19128 from Patent WO0170979.
ACCESSION   CQ412057
VERSION     CQ412057.1  GI:41319838
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Lillie,J.
  TITLE     Genes, compositions, kits, and method for identification,
            assessment, prevention, and therapy of ovarian cancer
  JOURNAL   Patent: WO 0170979-A 19128 27-SEP-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .4065
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.44e-190       Length:         4065
Score:                  2198.00         Matches:        418
Percent Similarity:     99.76%          Conservative:   0
Best Local Similarity:  99.76%          Mismatches:     1
```

```
Query Match:            99.64%          Indels:         0
DB:                     6               Gaps:           0

US-10-798-532-2 (1-419) x CQ412057 (1-4065)

Qy         1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
Db       268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy        21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db       328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy        41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
Db       388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy        61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
Db       448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy        81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db       508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy       101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
Db       568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy       121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
Db       628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy       141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
Db       688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy       161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
Db       748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy       181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
Db       808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy       201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
Db       868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy       221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
Db       928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy       241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
Db       988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy       261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
Db      1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy       281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
Db      1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy       301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
Db      1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy       321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
Db      1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287
```

```
Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1524

RESULT 6
AX188296
LOCUS       AX188296                4065 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 3991 from Patent WO0142467.
ACCESSION   AX188296
VERSION     AX188296.1  GI:15139769
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
  TITLE     Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
  JOURNAL   Patent: WO 0142467-A 3991 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .4065
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.44e-190      Length:         4065
Score:                  2198.00        Matches:        418
Percent Similarity:     99.76%         Conservative:   0
Best Local Similarity:  99.76%         Mismatches:     1
Query Match:            99.64%         Indels:         0
DB:                     6              Gaps:           0

US-10-798-532-2 (1-419) x AX188296 (1-4065)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1524

RESULT 7
AX166543
LOCUS       AX166543                1260 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 34 from Patent WO0138503.
ACCESSION   AX166543
VERSION     AX166543.1  GI:14546888
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
```

```
  AUTHORS   Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
            Flanagan,P. and Clary,D.S.
  TITLE     Novel human protein kinases and protein kinase-like enzymes
  JOURNAL   Patent: WO 0138503-A 34 31-MAY-2001;
            Sugen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1260
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.23e-190       Length:         1260
Score:                  2191.00         Matches:        417
Percent Similarity:     99.52%          Conservative:   0
Best Local Similarity:  99.52%          Mismatches:     2
Query Match:            99.32%          Indels:         0
DB:                     6               Gaps:           0

US-10-798-532-2 (1-419) x AX166543 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257

RESULT 8
AX680136
LOCUS       AX680136                1824 bp    DNA     linear   PAT 29-MAR-2003
DEFINITION  Sequence 21 from Patent WO0181555.
ACCESSION   AX680136
VERSION     AX680136.1  GI:29369916
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Gururajan,R., Patterson,C., Recipon,S.A., Tribouley,C.M.,
            Hafalia,A., Khan,F.Y., Yue,H., Au-Young,J., Bandman,O.,
            Baughn,M.R., Borowsky,M.L., Burford,N., Burrill,J.D., Elliot,V.S.,
            Gandhi,A.R., Kearney,L., Lal,P., Lu,D.A., Lu,Y., Tang,Y.T.,
            Zingler,K.A., Griffin,J.A., Hillman,J.J., Marcus,G.A., Nguyen,D.B.,
            Policky,J., Ramkumar,J., Thornton,M., Walia,N.K. and Walsh,R.T.
  TITLE     Human kinases
  JOURNAL   Patent: WO 0181555-A 21 01-NOV-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1824
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 3107952CB1"
ORIGIN

Alignment Scores:
Pred. No.:              2.05e-190       Length:         1824
Score:                  2191.00         Matches:        417
Percent Similarity:     99.52%          Conservative:   0
Best Local Similarity:  99.52%          Mismatches:     2
Query Match:            99.32%          Indels:         0
DB:                     6               Gaps:           0

US-10-798-532-2 (1-419) x AX680136 (1-1824)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         169 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 228

Qy          21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         229 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 288

Qy          41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         289 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 348

Qy          61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         349 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 408

Qy          81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         409 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 468

Qy         101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         469 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 528

Qy         121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         529 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 588

Qy         141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         589 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 648

Qy         161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         649 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 708

Qy         181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         709 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 768

Qy         201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         769 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 828

Qy         221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         829 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 888

Qy         241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         889 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 948

Qy         261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         949 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1008

Qy         281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1009 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1068

Qy         301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1069 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1128

Qy         321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
                ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        1129 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1188

Qy         341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1189 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1248

Qy         361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
                ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        1249 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1308

Qy         381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1309 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1368

Qy         401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1369 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1425
```

RESULT 9
RNKISRNA

```
LOCUS       RNKISRNA                1633 bp    mRNA    linear   ROD 08-JUL-2002
DEFINITION  R.norvegicus mRNA for KIS protein.
ACCESSION   X98374
VERSION     X98374.1  GI:1403531
KEYWORDS    kis gene.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Maucuer,A., Ozon,S., Manceau,V., Gavet,O., Lawler,S., Curmi,P. and
            Sobel,A.
  TITLE     KIS is a protein kinase with an RNA recognition motif
  JOURNAL   J. Biol. Chem. 272 (37), 23151-23156 (1997)
  MEDLINE   97435279
   PUBMED   9287318
REFERENCE   2  (bases 1 to 1633)
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-1996) A. Maucuer, INSERM U440, 17 rue du Fer
            Moulin, 75005 Paris, FRANCE
FEATURES             Location/Qualifiers
     source          1. .1633
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /tissue_type="brain"
                     /clone_lib="7"
                     /dev_stage="neonate"
     gene            1. .1633
                     /gene="KIS"
     CDS             39. .1298
                     /gene="KIS"
                     /codon_start=1
                     /protein_id="CAA67021.1"
                     /db_xref="GI:1403532"
                     /db_xref="GOA:Q63285"
                     /db_xref="Swiss-Prot:Q63285"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              5.05e-190       Length:         1633
Score:                  2186.00         Matches:        415
Percent Similarity:     99.52%          Conservative:   2
Best Local Similarity:  99.05%          Mismatches:     2
Query Match:            99.09%          Indels:         0
DB:                     10              Gaps:           0

US-10-798-532-2 (1-419) x RNKISRNA (1-1633)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        39 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 98

Qy        21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40

Db        99 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 158

Qy        41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60

Db       159 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 218

Qy        61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80

Db       219 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 278

Qy        81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100

Db       279 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 338

Qy       101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120

Db       339 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 398

Qy       121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140

Db       399 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 458

Qy       141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160

Db       459 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 518

Qy       161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180

Db       519 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 578

Qy       181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200

Db       579 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 638

Qy       201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220

Db       639 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 698

Qy       221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240

Db       699 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 758

Qy       241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260

Db       759 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 818

Qy       261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280

Db       819 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 878

Qy       281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300

Db       879 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 938

Qy       301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320

Db       939 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 998

Qy       321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340

Db       999 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1058

Qy       341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360

Db      1059 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1118

Qy       361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

Db      1119 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1178

Qy       381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400

Db      1179 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1238

Qy       401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419

Db      1239 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 1295
```

```
RESULT 10
RNU70372
LOCUS       RNU70372               3244 bp    mRNA    linear   ROD 08-SEP-1999
DEFINITION  Rattus norvegicus PAM COOH-terminal interactor protein 2 mRNA,
            complete cds.
ACCESSION   U70372
VERSION     U70372.2  GI:5821767
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 3244)
  AUTHORS   Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
            and Eipper,B.A.
  TITLE     Novel proteins that interact with the COOH-terminal cytosolic
            routing determinants of an integral membrane peptide-processing
            enzyme
  JOURNAL   J. Biol. Chem. 271 (45), 28636-28640 (1996)
  MEDLINE   97067094
   PUBMED   8910496
REFERENCE   2  (bases 1 to 3244)
  AUTHORS   Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
            Eipper,B.A. and Mains,R.E.
  TITLE     The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
            Determinants of the Peptide Processing Enzyme Peptidylglycine
            alpha-Amidating Monooxygenase
  JOURNAL   J. Biol. Chem. 274 (1999) In press
REFERENCE   3  (bases 1 to 3244)
  AUTHORS   Alam,R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
            and Eipper,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
            School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
            21205, USA
REFERENCE   4  (bases 1 to 3244)
  AUTHORS   Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
            Eipper,B.A. and Mains,R.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
            School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
            21205, USA
  REMARK    Sequence update by submitter
COMMENT     On Sep 3, 1999 this sequence version replaced gi:1698778.
FEATURES             Location/Qualifiers
     source          1. .3244
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
     CDS             162. .1421
                     /function="serine protein kinase"
                     /note="P-CIP2"
                     /codon_start=1
                     /product="PAM COOH-terminal interactor protein 2"
                     /protein_id="AAC53031.2"
                     /db_xref="GI:5821768"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
```

```
                    MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

Alignment Scores:
Pred. No.:              1.31e-189       Length:         3244
Score:                  2186.00         Matches:        415
Percent Similarity:     99.52%          Conservative:   2
Best Local Similarity:  99.05%          Mismatches:     2
Query Match:            99.09%          Indels:         0
DB:                     10              Gaps:           0

US-10-798-532-2 (1-419) x RNU70372 (1-3244)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 221

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 281

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        282 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 341

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 401

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 461

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 521

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 581

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 641

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 701

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 761

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 821

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 881

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 941

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 1001

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 1061

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 1121

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||  ||||||:::||||||
Db       1122 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1181

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1182 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1241

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1242 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1301

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1302 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1361

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1362 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 1418

RESULT 11
BC058732
LOCUS       BC058732                2041 bp    mRNA    linear   ROD 30-JUN-2004
DEFINITION  Mus musculus kinase interacting with leukemia-associated gene
            (stathmin), mRNA (cDNA clone MGC:64770 IMAGE:6414877), complete
            cds.
ACCESSION   BC058732
VERSION     BC058732.1  GI:37194892
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2041)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2041)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
Web site: http://genome.uiowa.edu
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A., Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K., Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J., Casavant,T., Soares,M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754447.

FEATURES Location/Qualifiers
source 1. .2041
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:64770 IMAGE:6414877"
/tissue_type="Brain, pool of enriched mouse brain tissues from 13.5,14.5,16.5 and 17.5 dpc"
/clone_lib="NIH_BMAP_FR0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
gene 1. .2041
/gene="Kist"
/note="synonym: KIS"
/db_xref="LocusID:16589"
/db_xref="MGI:1341908"
CDS 154. .1413
/gene="Kist"
/codon_start=1
/product="kinase interacting with leukemia-associated gene (stathmin)"
/protein_id="AAH58732.1"
/db_xref="GI:37194893"
/db_xref="LocusID:16589"
/db_xref="MGI:1341908"
/translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
MFDGKFVVATFYPLSAYKRGYLYQTLL"

ORIGIN

Alignment Scores:

Pred. No.:	8.5e-190	Length:	2041
Score:	2185.00	Matches:	415
Percent Similarity:	99.28%	Conservative:	1
Best Local Similarity:	99.05%	Mismatches:	3
Query Match:	99.05%	Indels:	0
DB:	10	Gaps:	0

US-10-798-532-2 (1-419) x BC058732 (1-2041)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 213

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 273

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        274 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 333

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 393

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 453

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 513

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 573

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 633

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        634 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 693

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        694 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 753

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        754 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 813

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        814 TCAGCTGTTGATCTGTGGAGCCTCGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAG 873

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        874 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAT 933

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        934 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 993

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        994 ATCAAAAGCATGCTTCATGATGACCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGC 1053

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1054 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACT 1113

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              |||||||||||||||||||||||||||||||||||||||||||||||  ||||||:::||||||
Db       1114 CCAGTGCTCAGGCTCCTCAATGTGCTGGATGATGATTACCTTGAAAATGAAGATGAATAT 1173

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1174 GAAGATGTTGTAGAAGATGTGAAAGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1233

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1234 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1293

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1294 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1353

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1354 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 1410
```

RESULT 12
AC129141_1/c

```
WPCOMMENT
  Sequence split into 4 fragments  LOCUS AC129141 Accession AC129141
     Fragment Name      Begin     End
     AC129141_0             1     110000
     AC129141_1        100001     210000
     AC129141_2        200001     310000
     AC129141_3        300001     363206
Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH2

Alignment Scores:
Pred. No.:              1.18e-186      Length:         110000
Score:                  2177.00        Matches:        414
Percent Similarity:     99.28%         Conservative:   2
Best Local Similarity:  98.81%         Mismatches:     3
Query Match:            98.69%         Indels:         0
DB:                     2              Gaps:           0

US-10-798-532-2 (1-419) x AC129141_1 (1-110000)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db      70316 ATGGCCGGGTCCTGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 70257

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70256 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 70197

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db      70196 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 70137

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70136 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 70077

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70076 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 70017

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70016 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 69957

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              |||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69956 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATACAGCACTGTGCCAGA 69897

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69896 GATGTTCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAA 69837

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69836 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 69777

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69776 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 69717

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69716 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 69657

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69656 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 69597

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69596 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 69537

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      69536 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 69477

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69476 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 69417

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69416 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 69357

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||:::||||||
Db      69356 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 69297

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69296 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 69237

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69236 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 69177

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69176 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGATGGGAAGTTTGTTGTG 69117

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69116 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 69060

RESULT 13
MMPKKIS
LOCUS       MMPKKIS                 1254 bp    mRNA    linear   ROD 08-JUL-2002
DEFINITION  Mus musculus partial mRNA for KIS protein kinase (Kis gene).
ACCESSION   Y10725
VERSION     Y10725.2  GI:21726712
KEYWORDS    Kis gene; protein kinase.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1254)
  AUTHORS   Maucuer,A., Ozon,S., Manceau,V., Gavet,O., Lawler,S., Curmi,P. and
            Sobel,A.
  TITLE     KIS is a protein kinase with an RNA recognition motif
  JOURNAL   J. Biol. Chem. 272 (37), 23151-23156 (1997)
  MEDLINE   97435279
   PUBMED   9287318
REFERENCE   2
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-1997) Maucuer A., INSERM U440, U440, 17 Rue du
            Fer Moulin, Paris, 75005, FRANCE
  REMARK    revised by [4]
REFERENCE   3  (bases 1 to 1254)
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) A. Maucuer, INSERM U440, U440, 17 Rue du
            Fer Moulin, Paris, 75005, FRANCE
COMMENT     On Jul 10, 2002 this sequence version replaced gi:1806129.
            Overlaps with related sequences X82320, X10725.
FEATURES             Location/Qualifiers
     source          1. .1254
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="K22"
                     /tissue_type="brain"
                     /clone_lib="lambda gt10"
                     /dev_stage="neonate"
     gene            1. .1254
                     /gene="Kis"
     CDS             11. .>1254
```

/gene="Kis"
/codon_start=1
/product="KIS protein kinase"
/protein_id="CAA71714.2"
/db_xref="GI:21726713"
/db_xref="GOA:P97343"
/db_xref="Swiss-Prot:P97343"
/translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
MFDGKFVVATFYPLSAYKRGYL"

ORIGIN

Alignment Scores:

Pred. No.:	8.34e-188	Length:	1254
Score:	2160.00	Matches:	410
Percent Similarity:	99.28%	Conservative:	1
Best Local Similarity:	99.03%	Mismatches:	3
Query Match:	97.91%	Indels:	0
DB:	10	Gaps:	0

US-10-798-532-2 (1-419) x MMPKKIS (1-1254)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 70

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 130

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        131 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 190

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 250

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 310

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 370

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 430

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 490

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 550

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 610

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 670

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 TCAGCTGTTGATCTGTGGAGCCTCGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAG 730

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        731 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAT 790

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        791 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 850

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              |||||||||||||||||||||||||    ||||||||||||||||||||||||||||||
Db        851 ATCAAAAGCATGCTTCATGATGACCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGC 910

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        911 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACT 970

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        971 CCAGTGCTCAGGCTCCTCAATGTGCTGGATGATGATTACCTTGAAAATGAAGATGAATAT 1030

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1031 GAAGATGTTGTAGAAGATGTGAAAGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1090

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1150

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1151 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1210

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeu 414
              |||||||||||||||||||||||||||||||||||||||||||
Db       1211 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTT 1252
```

RESULT 14
BC026046

LOCUS	BC026046 1367 bp mRNA linear PRI 30-JUN-2004
DEFINITION	Homo sapiens kinase interacting with leukemia-associated gene (stathmin), mRNA (cDNA clone MGC:39391 IMAGE:4649261), complete cds.
ACCESSION	BC026046
VERSION	BC026046.1 GI:19684094
KEYWORDS	MGC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1367)
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

```
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1367)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 42 Row: f Column: 6
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 21389438.
FEATURES             Location/Qualifiers
     source          1. .1367
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:39391 IMAGE:4649261"
                     /tissue_type="Eye, retinoblastoma"
                     /clone_lib="NIH_MGC_16"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1. .1367
                     /gene="KIS"
                     /db_xref="LocusID:127933"
     CDS             159. .1193
                     /gene="KIS"
                     /codon_start=1
                     /product="kinase interacting stathmin"
                     /protein_id="AAH26046.1"
                     /db_xref="GI:19684095"
                     /db_xref="LocusID:127933"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSATIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPPFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEGLC"
ORIGIN

Alignment Scores:
Pred. No.:              1.77e-172       Length:         1367
Score:                  1993.00         Matches:        386
Percent Similarity:     92.14%          Conservative:   1
Best Local Similarity:  91.90%          Mismatches:     2
Query Match:            90.34%          Indels:         31
DB:                     9               Gaps:           1

US-10-798-532-2 (1-419) x BC026046 (1-1367)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 218

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 278

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 338

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 398

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 458

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 518

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 578

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 638

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 698

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 758

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 818

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 878

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTACTATTGATCAC 938

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 998

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1058

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1118

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1119 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1178

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360

Db       1178 ------------------------------------------------------------ 1178

Qy        361 LeuValProLysGluAsnProGlyArgGlyGln-ValPheValGluTyrAlaAsnAlaGl 380
                                                  ::: ||||||||||||||||||||||||||
Db       1179 -----------------------------GAAGGTCTTTGTTGAGTATGCAAATGCTGG 1208
```

```
Qy        380 yAspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVa 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1209 TGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGT 1268

Qy        400 lAlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 GGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1326

RESULT 15
HSM805327
LOCUS       HSM805327               2940 bp    mRNA    linear   PRI 12-JUL-2002
DEFINITION  Homo sapiens mRNA; cDNA DKFZp667I0511 (from clone DKFZp667I0511).
ACCESSION   AL834136
VERSION     AL834136.1  GI:21739600
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2940)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp667I0511) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .2940
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:DKFZp667I0511"
                     /db_xref="taxon:9606"
                     /map="1"
                     /clone="DKFZp667I0511"
                     /tissue_type="lymph node"
                     /clone_lib="667 (synonym: hlno2). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     polyA_site      2923
ORIGIN

Alignment Scores:
Pred. No.:              6.49e-163       Length:         2940
Score:                  1893.50         Matches:        389
Percent Similarity:     63.46%          Conservative:   0
Best Local Similarity:  63.46%          Mismatches:     2
Query Match:            85.83%          Indels:         223
DB:                     9               Gaps:           1

US-10-798-532-2 (1-419) x HSM805327 (1-2940)

Qy         29 LeuGlySerGlySerSerAlaSerValTyrArgValArgCysCysGlyAsnProGlySer 48
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTTCGCTGCTGCGGCAACCCTGGCTCG 60

Qy         49 ProProGlyAlaLeuLysGlnPheLeuProProGlyThrThrGlyAlaAlaAlaSerAla 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGAACCACCGGGGCTGCGGCCTCTGCC 120

Qy         69 AlaGluTyrGlyPheArgLysGluArgAlaAlaLeuGluGlnLeuGlnGlyHisArgAsn 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTGGAACAGTTGCAGGGTCACAGAAAC 180

Qy         89 IleValThrLeuTyrGlyValPheThrIleHisPheSerProAsnValProSerArgCys 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCGTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGT 240

Qy        109 LeuLeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGln 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAG 300

Qy        129 GlyCysSerMetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPhe 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTGTTCCATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTT 360

Qy        149 LeuHisHisGluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAla 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTCATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCA 420

Qy        169 GluAsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGln--- 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGAATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCA-GGT 479

Qy        187 ------------------------------------------------------------ 187

Db        480 AAGAAATAACCTTTTCTTTTCTCTCGCAGTTTAAAATGTAGTGGTTGAAGCCAAATATTT 539

Qy        187 ------------------------------------------------------------ 187

Db        540 TTAAATGACAAGGGTAGTTTAGTAGGGGCTTACAAGTCATTTTAACCTAAATTACTGAAA 599

Qy        187 ------------------------------------------------------------ 187

Db        600 TGCCAGGGGATGCAGTAAATCAAAAGAATTTCAGTGCCACCGTGTAAAAAAAGGACATAA 659

Qy        187 ------------------------------------------------------------ 187

Db        660 TGTTTGGGTCAGGTAGGAGGAGCCAGTGAAGCATGAGAGCTCTGGAATTATTAAAAGTAC 719

Qy        187 ------------------------------------------------------------ 187

Db        720 TTAACCTACATAGAATGGTTTTTTAAATTTTATGCGTGTTTGGCTAATTGCTGAGTTACT 779

Qy        187 ------------------------------------------------------------ 187

Db        780 ATTCCTGGTCTCTTCAGTGTTGGGGCTGTATCTCGTGGCAGAATGCATATGATAGTTCTG 839

Qy        187 ------------------------------------------------------------ 187

Db        840 CCAAATCCTCTTTTCTCCATGGTTTATCATTTCTTTGCAGTGTCCTTATTATCAGCAGTG 899

Qy        187 ------------------------------------------------------------ 187

Db        900 CCTAATAGGTAAGAGACAATCCAAAGAAGACTTCTTTATGAGCGCTATATGGAAATGTAC 959

Qy        187 ------------------------------------------------------------ 187

Db        960 AGCTTGCCTCTGAGAACTGCCCAGTTAATGAACCAAAGAATTATGGTAATCAAGGACTAA 1019

Qy        187 ------------------------------------------------------------ 187

Db       1020 AAATTTTTAGCATTTAGCTTGGCAGTGGGTTTACTGCACTCTTAGGGAAGTAAATTCACT 1079

Qy        187 ------------------------------------------------------------ 187

Db       1080 GTAGAGATGGATAGAAAGGGAGCAGCTTTTTTATTGGTTGTAATATTTACTCATCTTTTT 1139

Qy        188 ---AspValLysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsn 206
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 TAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAAT 1199

Qy        207 CysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrp 226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 TGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGG 1259

Qy        227 SerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArg 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 AGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGA 1319
```

```
Qy        247 SerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAla 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCA 1379

Qy        267 ValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHis 286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 GTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCAT 1439

Qy        287 AspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIle 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 GATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATT 1499

Qy        307 ProPheAlaProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeu 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 CCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTG 1559

Qy        327 AsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAsp 346
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db       1560 AATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATATGAAGATGTTGTAGAAGAT 1619

Qy        347 ValLysGluGluCysGlnLysTyrGlyProValValSerLeuLeuValProLysGluAsn 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 GTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAAT 1679

Qy        367 ProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGlyAspSerLysAlaAlaGln 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 CCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAG 1739

Qy        387 LysLeuLeuThrGlyArgMetPheAspGlyLysPheValValAlaThrPheTyrProLeu 406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTG 1799

Qy        407 SerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||
Db       1800 AGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1838
```

Search completed: November 29, 2004, 18:10:34
Job time : 4620 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 502 Seconds
(without alignments)
4381.489 Million cell updates/sec

Title: US-10-798-532-2
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAFG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10798532/runat_29112004_112452_15882/app_query.fasta_1.1166
-DB=N_Geneseq_23Sep04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532_@CGN_1_1_886_@runat_29112004_112452_15882 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2206	100.0	1260	3	AAZ51355	Aaz51355 Human wil
2	2203	99.9	1260	3	AAZ51356	Aaz51356 Human tra
3	2198	99.6	1260	4	AAI66829	Aai66829 Human pro
4	2198	99.6	2008	6	ABQ99383	Abq99383 Human cod
5	2198	99.6	2622	4	AAI66828	Aai66828 Human pro
6	2198	99.6	4065	4	AAH72714	Aah72714 Human cer
7	2198	99.6	4065	5	ADL45238	Adl45238 Human ova
8	2191	99.3	1260	4	AAS06734	Aas06734 Polynucle
9	2191	99.3	1824	4	AAD18818	Aad18818 Human kin
10	2186	99.1	3244	10	ADB52857	Adb52857 Primary r
11	2045.5	92.7	3998	12	ADQ24401	Adq24401 Human sof
12	789	35.8	529	4	AAH70764	Aah70764 Human cer
13	766.5	34.7	722	3	AAA02535	Aaa02535 Human col
14	729	33.0	470	4	AAH71966	Aah71966 Human cer
15	636	28.8	461	4	AAH69101	Aah69101 Human cer
16	559.5	25.4	411	6	ABQ99225	Abq99225 Human ORF
17	361	16.4	231	3	AAA01259	Aaa01259 Human col
18	311.5	14.1	419	5	ADI73740	Adi73740 Human ova
19	311.5	14.1	419	5	ADI67345	Adi67345 Human ova
c 20	287	13.0	546	6	ABQ53127	Abq53127 Oligonucl
21	287	13.0	546	6	ABQ53126	Abq53126 Oligonucl
c 22	287	13.0	547	6	ABQ44465	Abq44465 Oligonucl
23	287	13.0	547	6	ABQ44464	Abq44464 Oligonucl
24	258	11.7	546	6	ABQ53129	Abq53129 Oligonucl
c 25	258	11.7	546	6	ABQ53128	Abq53128 Oligonucl
26	258	11.7	547	6	ABQ44463	Abq44463 Oligonucl
c 27	258	11.7	547	6	ABQ44462	Abq44462 Oligonucl
28	252	11.4	1065	10	ADD21399	Add21399 Golden ha
29	250	11.3	442	9	ACH28557	Ach28557 Human adu
30	249.5	11.3	1400	10	ADD21407	Add21407 Rat cycli
31	248	11.2	1708	10	ADD21395	Add21395 Mouse cyc
32	243.5	11.0	2472	3	AAC45716	Aac45716 Arabidops
33	243	11.0	1297	10	ADD21393	Add21393 Human cyc
34	243	11.0	2213	6	ABK84073	Abk84073 Human cDN
35	243	11.0	2213	10	ADD21403	Add21403 Chinese h
36	243	11.0	2213	11	ADI31711	Adi31711 Human cDN
37	243	11.0	2213	12	ADH17138	Adh17138 Human cdc
38	243	11.0	2328	10	ACC72837	Acc72837 Human can
39	243	11.0	2734	12	ADQ25034	Adq25034 Human sof
40	240	10.9	1209	10	ADD21401	Add21401 Golden ha
41	239.5	10.9	1635	2	AAV71073	Aav71073 CDK2-gree
42	239	10.8	1297	6	ABQ60787	Abq60787 Human 10C
43	239	10.8	1297	10	ADD21391	Add21391 Human cyc
44	238.5	10.8	1650	12	ADI45326	Adi45326 Rice isop
45	237.5	10.8	1544	10	ADD21409	Add21409 Rat cycli

ALIGNMENTS

RESULT 1
AAZ51355
ID AAZ51355 standard; DNA; 1260 BP.
XX
AC AAZ51355;
XX
DT 06-JUN-2000 (first entry)
XX
DE Human wild type serine/threonine kinase KIS (hKIS) gene.
XX
KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
KW dominant gene; cyclin-dependent kinase inhibitor; CKI; p27; modulator;
KW treatment; cell proliferative disease; vascular disorder; gene therapy;
KW atherosclerosis; restenosis; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1260
FT /*tag= a
FT /product= "Human wild type KIS protein"
FT /transl_except= (pos:556. .561, Asn)
XX
PN WO200011165-A1.
XX
PD 02-MAR-2000.
XX
PF 20-AUG-1999; 99WO-US018903.
XX
PR 21-AUG-1998; 98US-0097710P.

XX
PA (NABE/) NABEL G J.
PA (NABE/) NABEL E G.
XX
PI Nabel GJ, Nabel EG;
XX
DR WPI; 2000-237648/20.
DR P-PSDB; AAY70305.
XX
PT Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT proliferation.
XX
PS Claim 1; Page 56; 70pp; English.
XX
CC The present DNA sequence encodes a wild type human KIS (hKIS), isolated
CC from a human B-cell library, using a yeast two hybrid screening system.
CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
CC that controls cell proliferation and is localised predominantly in the
CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
CC The hKIS sequences are used to modulate cell proliferation and treat cell
CC proliferative and vascular diseases. The polynucleotide sequence may be
CC used in gene therapy to treat vascular disorders such as restenosis or
CC atherosclerosis
XX
SQ Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.75e-209	Length:	1260
Score:	2206.00	Matches:	419
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-10-798-532-2 (1-419) x AAZ51355 (1-1260)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257
```

RESULT 2
AAZ51356
ID AAZ51356 standard; DNA; 1260 BP.
XX
AC AAZ51356;
XX
DT 06-JUN-2000 (first entry)
XX
DE Human transdominant mutant serine/threonine kinase KIS (hKIS) gene.
XX
KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
KW transdominant gene; mutant; cyclin-dependent kinase inhibitor; CKI; p27;
KW modulator; treatment; cell proliferative disease; vascular disorder;
KW gene therapy; atherosclerosis; restenosis; ds.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	CDS	1. .1260
FT		/*tag= a
FT		/product= "Human transdominant mutant KIS protein"
FT		/transl_except= (pos:556. .561, Asn)
FT	misc_feature	161

```
FT                   /*tag=  b
FT                   /note= "A to G substitution resulting in a Lys to Arg
FT                   change in the amino acid sequence"
XX
PN   WO200011165-A1.
XX
PD   02-MAR-2000.
XX
PF   20-AUG-1999;   99WO-US018903.
XX
PR   21-AUG-1998;   98US-0097710P.
XX
PA   (NABE/) NABEL G J.
PA   (NABE/) NABEL E G.
XX
PI   Nabel GJ,  Nabel EG;
XX
DR   WPI; 2000-237648/20.
DR   P-PSDB; AAY70306.
XX
PT   Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT   for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT   proliferation.
XX
PS   Claim 7; Page 59; 70pp; English.
XX
CC   The present DNA sequence encodes a transdominant mutant human KIS (hKIS),
CC   constructed by site directed mutagenesis. A single nucleotide
CC   substitution (A to G) results in a lysine to arginine change in the
CC   protein sequence. hKIS is a serine/threonine kinase, that acts as an
CC   inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
CC   controls cell proliferation and is localised predominantly in the
CC   nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC   G1 phase. The hKIS sequences are used to modulate cell proliferation and
CC   treat cell proliferative and vascular diseases. The polynucleotide
CC   sequence may be used in gene therapy to treat vascular disorders such as
CC   restenosis or atherosclerosis
XX
SQ   Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                9.44e-209      Length:          1260
Score:                    2203.00        Matches:         418
Percent Similarity:       100.00%        Conservative:    1
Best Local Similarity:    99.76%         Mismatches:      0
Query Match:              99.86%         Indels:          0
DB:                       3              Gaps:            0

US-10-798-532-2 (1-419) x AAZ51356 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257

RESULT 3
AAI66829
ID   AAI66829 standard; cDNA; 1260 BP.
XX
AC   AAI66829;
XX
DT   07-JAN-2002  (first entry)
XX
DE   Human protein kinase polypeptide 13887 coding sequence.
XX
KW   Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KW   cytostatic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;
```

```
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1260
FT                   /*tag=  a
FT                   /product= "protein kinase 13887"
XX
PN   WO200173050-A2.
XX
PD   04-OCT-2001.
XX
PF   23-MAR-2001; 2001WO-US009483.
XX
PR   24-MAR-2000; 2000US-0191846P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Meyers R;
XX
DR   WPI; 2001-611632/70.
DR   P-PSDB; AAG65767.
XX
PT   New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT   useful in diagnosis of cancer or cellular proliferation or
PT   differentiation disorders and to screen for polypeptide modulators useful
PT   to treat such conditions.
XX
PS   Claim 1; Fig 10A-C; 169pp; English.
XX
CC   The invention provides novel human protein kinase polypeptides, 3714,
CC   16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC   protein kinase polypeptides can be expressed by standard recombinant
CC   methodology.3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC   are useful for diagnostic and screening methods to identify subjects (at
CC   risk of) having cancer or cellular proliferation and/or differentiation
CC   disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC   modulators are useful for the treatment of cancer, particularly colon
CC   cancer or cellular proliferation and/or differentiation disorders. Other
CC   disorders associated with 3714, 16742, 23546 or 13887 expression or
CC   activity that can be treated include bone related disorders, inflammatory
CC   disorders, autoimmune diseases, cardiovascular disorders and liver
CC   diseases. The present sequence represents a human protein kinase
CC   polypeptide 13887 coding sequence
XX
SQ   Sequence 1260 BP; 312 A; 279 C; 334 G; 335 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.96e-208      Length:          1260
Score:                    2198.00        Matches:         418
Percent Similarity:       99.76%         Conservative:    0
Best Local Similarity:    99.76%         Mismatches:      1
Query Match:              99.64%         Indels:          0
DB:                       4              Gaps:            0

US-10-798-532-2 (1-419) x AAI66829 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257

RESULT 4
ABQ99383
ID   ABQ99383 standard; cDNA; 2008 BP.
XX
```

AC ABQ99383;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human coding sequence SEQ ID 116.
XX
KW Human; expressed sequence tag; EST; haematopoietic disorder;
KW central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200259260-A2.
XX
PD 01-AUG-2002.
XX
PF 16-NOV-2001; 2001WO-US042950.
XX
PR 17-NOV-2000; 2000US-00714936.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-590824/63.
DR N-PSDB; ABP64797.
XX
PT New isolated polynucleotide, useful in research, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
PT aberrant protein expression or biological activity.
XX
PS Claim 1; SEQ ID NO 116; 394pp; English.
XX
CC The present invention relates to novel human coding sequences (ABQ99268-
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC therapeutic, diagnostic and research methods. The polynucleotides may be
CC used in the field of molecular biology as hybridisation probes, primers
CC for PCR, for chromosome and gene mapping, for the recombinant production
CC of protein, or in generation of anti-sense DNA or RNA. The
CC polynucleotides are useful in diagnostics as expressed sequence tags
CC (ESTs) for identifying expressed genes or for physical mapping of the
CC human genome. The proteins may be used as molecular weight markers, or as
CC nutritional sources or supplements. The proteins may be used to maintain
CC and expand cell population in a totipotential or pluripotential state
CC useful for re-engineering damaged or diseased tissues, transplantation,
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC polynucleotides and proteins are useful for preventing, treating or
CC ameliorating disorders involving aberrant protein expression or
CC biological activity, e.g. haematopoietic disorders, central/peripheral
CC nervous system diseases, mechanical and traumatic disorders, non-healing
CC wounds, immune deficiencies and disorders, infectious diseases caused by
CC viral, bacterial or fungal infection, autoimmune disorders, allergic
CC reactions and conditions, coagulation disorders, or cancer. The
CC polynucleotide sequences of the invention were assembled from ESTs
CC isolated mainly by sequencing by hybridisation, and in some cases,
CC sequences obtained from one or more public databases. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.86e-208	Length:	2008
Score:	2198.00	Matches:	418
Percent Similarity:	99.76%	Conservative:	0
Best Local Similarity:	99.76%	Mismatches:	1
Query Match:	99.64%	Indels:	0
DB:	6	Gaps:	0

US-10-798-532-2 (1-419) x ABQ99383 (1-2008)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 138

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 198

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 258

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 318

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 378

Qy        101 SerProAsnValProSerArgCysLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 438

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 498

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 558

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 618

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 678

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 738

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 798

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 858

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 918

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 978

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        979 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1038

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1039 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1098
```

```
Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1099 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1158

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1159 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1218

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1219 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1278

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1279 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1335
```

RESULT 5
AAI66828

```
ID   AAI66828 standard; cDNA; 2622 BP.
XX
AC   AAI66828;
XX
DT   07-JAN-2002  (first entry)
XX
DE   Human protein kinase polypeptide 13887 encoding cDNA.
XX
KW   Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KW   cytostatic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1. .266
FT                   /*tag=  a
FT   CDS             267. .1529
FT                   /*tag=  b
FT                   /product= "protein kinase 13887"
FT                   /note= "coding sequence specifically claimed"
FT   3'UTR           1530. .2622
FT                   /*tag=  c
XX
PN   WO200173050-A2.
XX
PD   04-OCT-2001.
XX
PF   23-MAR-2001; 2001WO-US009483.
XX
PR   24-MAR-2000; 2000US-0191846P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Meyers R;
XX
DR   WPI; 2001-611632/70.
DR   P-PSDB; AAG65767.
XX
PT   New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT   useful in diagnosis of cancer or cellular proliferation or
PT   differentiation disorders and to screen for polypeptide modulators useful
PT   to treat such conditions.
XX
PS   Claim 1; Fig 10A-C; 169pp; English.
XX
CC   The invention provides novel human protein kinase polypeptides, 3714,
CC   16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC   protein kinase polypeptides can be expressed by standard recombinant
CC   methodology.3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC   are useful for diagnostic and screening methods to identify subjects (at
CC   risk of) having cancer or cellular proliferation and/or differentiation
CC   disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC   modulators are useful for the treatment of cancer, particularly colon
CC   cancer or cellular proliferation and/or differentiation disorders. Other
CC   disorders associated with 3714, 16742, 23546 or 13887 expression or
CC   activity that can be treated include bone related disorders, inflammatory
CC   disorders, autoimmune diseases, cardiovascular disorders and liver
CC   diseases. The present sequence represents a human protein kinase
CC   polypeptide 13887 encoding cDNA
XX
SQ   Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	8.67e-208	Length:	2622
Score:	2198.00	Matches:	418
Percent Similarity:	99.76%	Conservative:	0
Best Local Similarity:	99.76%	Mismatches:	1
Query Match:	99.64%	Indels:	0
DB:	4	Gaps:	0

US-10-798-532-2 (1-419) x AAI66828 (1-2622)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 326

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 386

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 446

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 506

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 566

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 626

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 686

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 746

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 806

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 866

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 926

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 986

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1046

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1106
```

```
Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1166

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1226

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db       1227 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1286

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1346

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1406

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1466

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1523
```

RESULT 6
AAH72714
ID AAH72714 standard; cDNA; 4065 BP.
XX
AC AAH72714;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 3988.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 810-811; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	1.65e-207	Length:	4065
Score:	2198.00	Matches:	418
Percent Similarity:	99.76%	Conservative:	0
Best Local Similarity:	99.76%	Mismatches:	1
Query Match:	99.64%	Indels:	0
DB:	4	Gaps:	0

US-10-798-532-2 (1-419) x AAH72714 (1-4065)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1524
```

RESULT 7
ADL45238
ID ADL45238 standard; DNA; 4065 BP.
XX
AC ADL45238;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human ovarian cancer DNA marker #19128.
XX
KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
OS Homo sapiens.
XX
PN WO200170979-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009126.
XX
PR 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lee J, Lillie J;
XX
DR WPI; 2001-611502/70.
XX
PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX
PS Disclosure; SEQ ID NO 19128; 106pp; English.
XX
CC The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.
XX
SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	1.65e-207	Length:	4065
Score:	2198.00	Matches:	418
Percent Similarity:	99.76%	Conservative:	0
Best Local Similarity:	99.76%	Mismatches:	1
Query Match:	99.64%	Indels:	0
DB:	5	Gaps:	0

US-10-798-532-2 (1-419) x ADL45238 (1-4065)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
```

```
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220

Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240

Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260

Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280

Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300

Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320

Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340

Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360

Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400

Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419

Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1524
```

RESULT 8
AAS06734
ID AAS06734 standard; cDNA; 1260 BP.
XX
AC AAS06734;
XX
DT 12-SEP-2001 (first entry)
XX
DE Polynucleotide sequence encoding human protein kinase #34.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US032085.
XX
PR 24-NOV-1999; 99US-0167482P.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX
DR WPI; 2001-343950/36.
DR P-PSDB; AAU03534.
XX
PT Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections.
XX
PS Example 1; Fig 1; 433pp; English.
XX
CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
CC protein kinases have been identified as members of the tyrosine or
CC serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be used for
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC polypeptides may be used as antigens in the production of antibodies
CC against the protein kinases and in assays to identify modulators of
CC protein kinase expression and activity
XX
SQ Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.47e-207	Length:	1260
Score:	2191.00	Matches:	417
Percent Similarity:	99.52%	Conservative:	0
Best Local Similarity:	99.52%	Mismatches:	2
Query Match:	99.32%	Indels:	0
DB:	4	Gaps:	0

US-10-798-532-2 (1-419) x AAS06734 (1-1260)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20

Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40

Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60

Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80

Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100

Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120

Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140

Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
```

```
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257
```

```
RESULT 9
AAD18818
ID   AAD18818 standard; cDNA; 1824 BP.
XX
AC   AAD18818;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human kinase (PKIN)-3 cDNA.
XX
KW   Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
KW   cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
KW   AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
KW   atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
KW   myasthenia gravis; cirrhosis; cataract; growth and development disorder;
KW   seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW   lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW   asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
KW   antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             169. .1428
FT                   /*tag=  a
FT                   /product= "Human PKIN-3 protein"
XX
PN   WO200181555-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-US012992.
XX
PR   20-APR-2000; 2000US-0199021P.
PR   28-APR-2000; 2000US-0200226P.
PR   05-MAY-2000; 2000US-0202339P.
PR   11-MAY-2000; 2000US-0203505P.
PR   18-MAY-2000; 2000US-0205564P.
PR   26-MAY-2000; 2000US-0207739P.
PR   01-JUN-2000; 2000US-0208795P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Gandhi AR,  Tribouley CM,  Kearney L,  Griffin JA,  Nguyen DB;
PI   Bandman O,  Lu DAM,  Lal P,  Burford N,  Khan FA,  Walia NK,  Yao MG;
PI   Patterson C,  Burrill JD,  Marcus GA,  Zingler KA,  Recipon SA,  Lu Y;
PI   Policky JL,  Thornton M,  Tang YT,  Hafalia A,  Elliott VS,  Baughn MR;
PI   Walsh RT,  Ramkumar J,  Borowsky ML,  Au-Young J,  Hillman JL;
PI   Gururajan R;
XX
DR   WPI; 2001-611740/70.
DR   P-PSDB; AAE11769.
XX
PT   Human kinases and nucleic acids, useful for preventing diagnosing and
PT   treating cancers, inflammation and immune disorders.
XX
PS   Claim 5; Page 152; 166pp; English.
XX
CC   The present invention relates to human kinases (PKIN) and the nucleic
CC   acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
CC   used in the prevention, diagnosis and treatment of diseases cancers,
CC   adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
CC   acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
CC   gout, microbial infections, cardiovascular disease and/or inflammation,
CC   myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
CC   infarction, cataract, growth and development disorder, seizure disorder,
CC   pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
CC   disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
CC   PKIN may be used to treat disorders associated with decreased PKIN
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of PKIN by expressing inactive proteins or to
CC   supplement the patients own production of PKIN. PKIN nucleic acids may be
CC   used to produce the PKIN polypeptide, by inserting the nucleic acids into
CC   a host cell and culturing the cell to express the protein. PKIN nucleic
CC   acid and its complementary sequences may also be used as DNA probes in
CC   diagnostic assays to detect and quantitate the presence of similar
CC   nucleic acid sequences in samples and therefore which patients may be in
CC   need of restorative therapy. The present sequence is human PKIN-3 cDNA
XX
SQ   Sequence 1824 BP; 434 A; 416 C; 472 G; 502 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:              2.53e-207       Length:         1824
Score:                  2191.00         Matches:        417
Percent Similarity:     99.52%          Conservative:   0
Best Local Similarity:  99.52%          Mismatches:     2
Query Match:            99.32%          Indels:         0
DB:                     4               Gaps:           0

US-10-798-532-2 (1-419) x AAD18818 (1-1824)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
```

```
Db        169 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 228

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40

Db        229 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 288

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60

Db        289 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 348

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80

Db        349 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 408

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100

Db        409 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 468

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120

Db        469 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 528

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140

Db        529 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 588

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160

Db        589 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 648

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180

Db        649 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 708

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200

Db        709 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 768

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220

Db        769 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 828

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240

Db        829 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 888

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260

Db        889 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 948

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280

Db        949 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1008

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300

Db       1009 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1068

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320

Db       1069 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1128

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340

Db       1129 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1188

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360

Db       1189 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1248

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

Db       1249 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1308

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400

Db       1309 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1368

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419

Db       1369 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1425
```

RESULT 10
ADB52857
ID ADB52857 standard; DNA; 3244 BP.
XX
AC ADB52857;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3399.
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN WO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
DR WPI; 2003-731472/69.
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 3399; 874pp; English.
XX
CC The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity

```
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 3244 BP; 800 A; 710 C; 819 G; 915 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.84e-206       Length:         3244
Score:                  2186.00         Matches:        415
Percent Similarity:     99.52%          Conservative:   2
Best Local Similarity:  99.05%          Mismatches:     2
Query Match:            99.09%          Indels:         0
DB:                     10              Gaps:           0

US-10-798-532-2 (1-419) x ADB52857 (1-3244)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 221

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 281

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              |||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        282 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 341

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 401

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 461

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 521

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 581

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 641

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 701

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 761

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 821

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 881

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 941

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 1001

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 1061

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 AGCCCATTCTTTAGCATTCCTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 1121

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||:::||||||
Db       1122 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1181

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1182 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1241

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1242 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1301

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1302 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1361

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1362 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 1418

RESULT 11
ADQ24401
ID   ADQ24401 standard; DNA; 3998 BP.
XX
AC   ADQ24401;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 7221.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 7221; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
```

```
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 3998 BP; 1066 A; 777 C; 885 G; 1255 T; 0 U; 15 Other;

Alignment Scores:
Pred. No.:              2.29e-192       Length:         3998
Score:                  2045.50         Matches:        396
Percent Similarity:     94.76%          Conservative:   2
Best Local Similarity:  94.29%          Mismatches:     21
Query Match:            92.72%          Indels:         1
DB:                     12              Gaps:           1

US-10-798-532-2 (1-419) x ADQ24401 (1-3998)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 219

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 279

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 339

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||
Db        340 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTAAGCATGTTTCTTACCAGACCG 399

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHis--- 99
              :::                          ||||||||||||||||||||||||||
Db        400 AAGGTCTGTGTTGACCTTAACCGGAGAGTGACTTTGTATGGAGTGTTTACAATCCAGGTT 459

Qy        100 PheSerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSer 119
                  :::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 GTTCTCCAAAGTGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCG 519

Qy        120 GluLeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAla 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 GAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCC 579

Qy        140 ArgAspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeu 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 CGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTC 639

Qy        160 LysProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGly 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 AAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGA 699

Qy        180 LeuSerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAla 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 CTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCT 759

Qy        200 ProGluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCys 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 CCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGT 819

Qy        220 ThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 879

Qy        240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 939

Qy        260 HisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAsp 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        940 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 999

Qy        280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 1059

Qy        300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCC 1119

Qy        320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGlu 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAA 1179

Qy        340 TyrGluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSer 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 TATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCT 1239

Qy        360 LeuLeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAla 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 CTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCT 1299

Qy        380 GlyAspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheVal 399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 GGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTT 1359

Qy        400 ValAlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1360 GTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1419

RESULT 12
AAH70764
ID   AAH70764 standard; cDNA; 529 BP.
XX
AC   AAH70764;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 2038.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer and
PT   for assessing and detecting compounds for treating the cancer.
XX
PS   Claim 1; Page 434; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
```

```
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              8.45e-69        Length:         529
Score:                  789.00          Matches:        159
Percent Similarity:     95.78%          Conservative:   0
Best Local Similarity:  95.78%          Mismatches:     5
Query Match:            35.77%          Indels:         2
DB:                     4               Gaps:           0

US-10-798-532-2 (1-419) x AAH70764 (1-529)

Qy        220 ThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 85

Qy        240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 145

Qy        260 HisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAsp 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 205

Qy        280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 265

Qy        300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCC 325

Qy        320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGlu 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAA 385

Qy        340 TyrGluAsp-ValValGluAspValLysGlu-GluCysGlnLysTyrGlyProValValS 359
              |||||||| |||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        386 TATGAAGATTGTTGTAGAAGATGTAAAAGAGGGAGTGTCAAAAATATGGACCAGGGGTAT 445

Qy        359 erLeuLeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnA 379
              ||||||||   ||||||||||||||||||||||   ||||||||||||||||||||||   |
Db        446 CTCTACTTGGTCCAAAGGAAAATCCTGGCAGAGGAACAGTCTTTGTTGAGTATGCAAAGG 505

Qy        379 laGlyAspSerLys 383
              |||||||||||||
Db        506 CTGGGGATTCAAAG 519

RESULT 13
AAA02535
ID   AAA02535 standard; cDNA; 722 BP.
XX
AC   AAA02535;
XX
DT   19-MAY-2000  (first entry)
XX
DE   Human colon cancer cell line polynucleotide sequence SEQ ID NO:2526.
XX
KW   Human; colon cancer; tumour; diagnosis; gene expression product; probe;
KW   detection; cancerous state; metastasis; identification; breast cancer;
KW   oestrogen receptor-positive breast cancer; therapy;
KW   oestrogen receptor-negative breast cancer; lung cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO9958675-A2.
XX
PD   18-NOV-1999.
XX
PF   13-MAY-1999;   99WO-US010602.
XX
PR   14-MAY-1998;   98US-0085426P.
PR   15-MAY-1998;   98US-0085537P.
PR   15-MAY-1998;   98US-0085696P.
PR   21-OCT-1998;   98US-0105234P.
PR   27-OCT-1998;   98US-0105877P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI   Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI   Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI   Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
DR   WPI; 2000-126369/11.
XX
PT   Polynucleotide library used to determine cancerous states of mammalian
PT   cells.
XX
PS   Claim 1; Page 1020; 1097pp; English.
XX
CC   AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC   libraries constructed from human colon cancer cell lines. The present
CC   invention also describes a method of detecting differentially expressed
CC   genes correlated with a cancerous state of a mammalian cell, comprising
CC   detecting at least one differentially expressed gene product in a test
CC   sample derived from a cell suspected of being cancerous, where detection
CC   of the differentially expressed gene product is correlated with a
CC   cancerous state of the cell from which the test sample was derived. The
CC   polynucleotides sequences can be used in a method for detecting
CC   differentially expressed genes correlated with a cancerous state of a
CC   mammalian cell. The polynucleotides can also be used as probes for
CC   detecting and mapping related genes. They can be used in diagnosis and
CC   prognosis of diseases and disorders (e.g. identification of pre-
CC   metastatic or metastatic cancerous states, stages of cancer, or
CC   responsiveness of cancer to therapy). This is particularly for breast
CC   cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC   negative breast cancer, lung cancer, and colon cancer
XX
SQ   Sequence 722 BP; 215 A; 129 C; 164 G; 193 T; 0 U; 21 Other;

Alignment Scores:
Pred. No.:              2.29e-66        Length:         722
Score:                  766.50          Matches:        182
Percent Similarity:     84.86%          Conservative:   3
Best Local Similarity:  83.49%          Mismatches:     20
Query Match:            34.75%          Indels:         14
DB:                     3               Gaps:           1

US-10-798-532-2 (1-419) x AAA02535 (1-722)

Qy        152 GluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGlu 171
              ||||||||||||||||||   ||||||   ||||||||||||||||||||||||||||||
Db          1 GAGGGCTATGTCCATGCGGNCCTCAAACNACGTAACATATTGTGGAGTGCAGAGAATGAA 60

Qy        172 CysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspValLysTyr 191
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db         61 TGTTTTAAACTCATTGACTTTGGACTTANCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 120

Qy        192 IleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAla 211
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||   |||
Db        121 ATTCAGACAGACGGGTATCGGGCTCCAGAACA-GAATTGCAAAATTGCTTGGCCCANGCT 179

Qy        212 GlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIle 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 239

Qy        232 LeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGluTrpLys 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 299
```

```
Qy        252 AlaAsnSer-SerAlaIleIle-AspHisIlePheAlaSerLysAla--ValValAsnAl 270
              ||||||||| |||||||||||| ||||||||||||||||||||:::     ||||||||
Db        300 GCAAACAGTTTCTGCTATTATTTGATCACATATTTGCCAGTAAAAGCAANTGGTGAATGC 359

Qy        270 aAlaIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAsp-ProS 290
              |||||||||||||||||||||   |||||||||||||||||||||||||||||| ||||
Db        360 CGCAATTCCAGCCTATCACCTAANAGACCTTATCAAAAGCATGCTTCATGATGATCCCAA 419

Qy        290 erArg-ArgIleProAlaGluMetAlaLeu-CysSerProPhePheSer-IleProPheA 309
              ||||| ||||||      |||||||||||| |||   |||||||||||| ||||||||||
Db        420 GCAGGAAGAATTTCTNCTGAAATGGCATTGGTGCANCCCATTCTTTAGCNATTCCTTTTG 479

Qy        309 laProHisIleGluAspLeuValMetLeuProThrProVal-LeuArgLeuLeuAsnVal 328
              |||||||||||||||||||   ||||||  |||||||| ||||||||||||||||||||
Db        480 CCCCTCATATTGAAGATCTGNTCATGCTTTCCACTCCAGTGGCTAAGACTGCTGAATGTG 539

Qy        329 Leu-AspAspAspTyrLeuGlyAsnGlu--GluGluTyrGluAspValValGluAspVal 347
              |||    ||||||||||||   ||||||  ::::::||||||||||||||||||||||
Db        540 CTGGGNTGATGATTATCTTGAGAATGAAAGAAGGATTATGAAGATGTTGTTGAAGATGNT 599

Qy        348 LysGluGluCysGlnLys---TyrGlyPro 356
              |||          |||      ||||||
Db        600 AAAAGAAGAAGTGGCAAAAATNTTGGACCA 629

RESULT 14
AAH71966
ID   AAH71966 standard; cDNA; 470 BP.
XX
AC   AAH71966;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 3240.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer and
PT   for assessing and detecting compounds for treating the cancer.
XX
PS   Claim 1; Page 636; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                6.51e-63      Length:          470
Score:                    729.00        Matches:         147
Percent Similarity:       96.71%        Conservative:    0
Best Local Similarity:    96.71%        Mismatches:      3
Query Match:              33.05%        Indels:          2
DB:                       4             Gaps:            0

US-10-798-532-2 (1-419) x AAH71966 (1-470)

Qy        220 ThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 77

Qy        240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 137

Qy        260 HisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAsp 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 197

Qy        280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 257

Qy        300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCC 317

Qy        320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGlu 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAA 377

Qy        340 TyrGluAsp-ValValGluAspValLysGlu-GluCysGlnLysTyrGlyProValValS 359
              |||||||| |||||||||||||||||||| ||||||||||||||||||||    ||||
Db        378 TATGAAGATTGTTGTAGAAGATGTAAAAGAGGGAGTGTCAAAAATATGGACCAGGGGTAT 437

Qy        359 erLeuLeuValProLysGluAsnProGlyArg 369
              ||||||||   ||||||||||||||||||||
Db        438 CTCTACTTGGTCCAAAGGAAAATCCTGGCAGA 469

RESULT 15
AAH69101
ID   AAH69101 standard; cDNA; 461 BP.
XX
AC   AAH69101;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 375.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
```

DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 168-169; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 461 BP; 134 A; 91 C; 109 G; 127 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.1e-53	Length:	461
Score:	636.00	Matches:	138
Percent Similarity:	93.33%	Conservative:	2
Best Local Similarity:	92.00%	Mismatches:	5
Query Match:	28.83%	Indels:	6
DB:	4	Gaps:	0

US-10-798-532-2 (1-419) x AAH69101 (1-461)

```
Qy        220 ThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 81

Qy        240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 141

Qy        260 HisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAsp 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 201

Qy        280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 261

Qy        300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
              |||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db        262 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATAT-GAAGATCTGGTCATGCTTCCC 320

Qy        320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGlu 339
              |||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db        321 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAA 380

Qy        340 TyrGluAsp--ValValGluAspValLysGlu--GluCysGlnLysTyrGlyPro-ValV 358
              |||||||||  ::::::||||||||||||||     ||||||||||||||||| ||||
Db        381 TATGAAGATTGTTGTTAGAAGATGTAAAAGAGGGAGGTGTCAAAAATATGGACCAGGTGG 440

Qy        358 alSerLeuLeuValProLys 364
              ||||||||||      |||
Db        441 TATCTCTACTTTGTTCCAAA 460
```

Search completed: November 29, 2004, 11:42:18
Job time : 521 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 3159.5 Seconds
(without alignments)
4832.489 Million cell updates/sec

Title: US-10-798-532-2
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAFG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10798532/runat_29112004_112453_15901/app_query.fasta_1.1166
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532_@CGN_1_1_6628_@runat_29112004_112453_15901 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2174	98.5	1763	3	AK013347	AK013347 Mus muscu
	2	1711	77.6	2846	3	AK030152	AK030152 Mus muscu
	3	1521.5	69.0	1079	4	BM468107	BM468107 AGENCOURT
	4	1274.5	57.8	1015	6	BY712344	BY712344
	5	1242	56.3	917	4	BM451184	BM451184 AGENCOURT
	6	1170	53.0	823	7	CO771206	CO771206 testis_ES
	7	1162	52.7	779	5	BU447831	603767571
	8	1159	52.5	896	7	CO773091	CO773091 testis_ES
	9	1133	51.4	781	7	CN232453	CN232453 WLB074D11
	10	1110	50.3	932	5	BU541306	BU541306 AGENCOURT
	11	1106	50.1	717	5	BU385761	603582650
	12	1091	49.5	899	2	BE895119	BE895119 601436041
	13	1079	48.9	772	4	BI562420	BI562420 603256316
	14	1075	48.7	794	5	BU205357	603104569
	15	1073.5	48.7	834	7	CF593363	CF593363 AGENCOURT
	16	1068.5	48.4	875	4	BI259568	BI259568 602968330
	17	1066	48.3	888	5	BQ225071	BQ225071 AGENCOURT
	18	1053	47.7	682	4	BI333581	BI333581 602997069
	19	1050	47.6	657	4	BG547675	BG547675 602575520
	20	1028	46.6	844	4	BG474616	BG474616 602517423
	21	1014	46.0	602	7	CO737403	CO737403 SlLT02c17
	22	1008.5	45.7	876	5	BU184346	BU184346 AGENCOURT
	23	990.5	44.9	906	4	BG386905	BG386905 602454704
	24	968.5	43.9	696	5	BU439278	BU439278 604147528
	25	967	43.8	631	5	BU112717	BU112717 603131421
	26	966.5	43.8	1003	5	BQ922020	BQ922020 AGENCOURT
	27	962	43.6	755	5	BU371858	BU371858 603568233
	28	958	43.4	699	5	BU707284	BU707284 UI-M-FR0-
	29	944	42.8	604	5	BX261107	BX261107 BX261107
	30	909.5	41.2	716	7	CF745774	CF745774 UI-M-GV0-
	31	900	40.8	570	5	BU293329	BU293329 604166812
	32	896	40.6	670	7	CN359673	CN359673 170004245
	33	892	40.4	928	5	BQ424438	BQ424438 AGENCOURT
	34	880.5	39.9	824	2	BE568528	BE568528 601342425
	35	870.5	39.5	820	7	CK597557	CK597557 AGENCOURT
	36	862	39.1	672	5	BX261106	BX261106 BX261106
	37	861.5	39.1	1063	6	CD510246	CD510246 CDA100-G0
	38	858	38.9	785	6	CB990192	CB990192 AGENCOURT
c	39	850	38.5	523	2	BF929410	BF929410 IL2-NT020
	40	828.5	37.6	1063	5	BU356643	BU356643 603474124
	41	818.5	37.1	743	4	BI562440	BI562440 603256324
	42	818	37.1	763	7	CK364431	CK364431 AGENCOURT
	43	801	36.3	933	2	BE886040	BE886040 601507192
c	44	796	36.1	986	5	BU751496	BU751496 CH4#002_E
	45	793	35.9	484	5	BX281211	BX281211 BX281211

ALIGNMENTS

RESULT 1
AK013347
LOCUS AK013347 1763 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810454004 product:kinase interacting with leukemia-associated gene (stathmin), full insert sequence.
ACCESSION AK013347
VERSION AK013347.1 GI:12850651
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,

Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1763)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.
FEATURES Location/Qualifiers
source 1. .1763
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2810454O04"
/db_xref="taxon:10090"
/clone="2810454O04"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10, 11 days embryo"
CDS 159. .1418
/note="unnamed protein product; kinase interacting with leukemia-associated gene (stathmin) (MGD|MGI:1341908) putative"
/codon_start=1
/protein_id="BAB28802.1"
/db_xref="GI:12850652"
/translation="MAGSGCAWGAEPPRFLEAPGRLWQVQSRLGSGSSASVYRVRCCG
TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALXQLQGHRNIVTLYGVFTIHFSP
NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
CTSAVDLWSLGITLLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

Alignment Scores:

Pred. No.:	2.59e-217	Length:	1763
Score:	2174.00	Matches:	413
Percent Similarity:	98.81%	Conservative:	1
Best Local Similarity:	98.57%	Mismatches:	5
Query Match:	98.55%	Indels:	0
DB:	3	Gaps:	0

US-10-798-532-2 (1-419) x AK013347 (1-1763)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 218

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 278

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 338

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 398

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 GANCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 458

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 518

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 578

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 638

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTAAGCTTATTGACTTTGGACTC 698

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 758

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 818

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        819 TCAGCTGTTGATCTGTGGAGCCTCGGAATCACTTTACTGGAAATGTTCTCAGGAATGAAG 878

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAT 938

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 998

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
                 ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        999 ATCAAAAGCATGCTTCATGATGACCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGC 1058

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACT 1118

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
                 ||||||||||||||||||||||||||||||||||||||||  ||||||:::||||||
Db       1119 CCAGTGCTCAGGCTCCTCAATGTGCTGGATGATGATTACCTTGAAAATGAAGATGAATAT 1178

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1179 GAAGATGTTGTAGAAGATGTGAAAGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1238

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1239 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1298

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1299 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1358

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1359 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 1415
```

RESULT 2
AK030152

```
LOCUS       AK030152               2846 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4933401D07 product:kinase interacting with
            leukemia-associated gene (stathmin), full insert sequence.
ACCESSION   AK030152
VERSION     AK030152.1  GI:26326136
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 2846)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.jp/
            URL:http://fantom.gsc.riken.jp/.
FEATURES             Location/Qualifiers
     source          1. .2846
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4933401D07"
                     /db_xref="taxon:10090"
                     /clone="4933401D07"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             238. .1230
                     /note="unnamed protein product; kinase interacting with
                     leukemia-associated gene (stathmin) (MGD|MGI:1341908)
                     putative"
                     /codon_start=1
                     /protein_id="BAC26812.1"
                     /db_xref="GI:26326137"
                     /translation="MTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMW
                     MIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY
                     IQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQE
                     WKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALCSPFFSIP
                     FAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSLLVPKE
                     NPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL"
     polyA_signal    2823. .2828
                     /note="putative"
     polyA_site      2846
                     /note="putative"
ORIGIN

Alignment Scores:
Pred. No.:              2.78e-168       Length:         2846
Score:                  1711.00         Matches:        326
Percent Similarity:     99.39%          Conservative:   2
```

```
Best Local Similarity:  98.79%          Mismatches:     2
Query Match:            77.56%          Indels:         0
DB:                     3               Gaps:           0

US-10-798-532-2 (1-419) x AK030152 (1-2846)

Qy          90 ValThrLeuTyrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
               :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         238 ATGACTTTATATGGAGTCTTTACCATACACTTCTCTCCCAATGTGCCATCACGCTGTCTG 297

Qy         110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGlnGly 129
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         298 CTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTTTTATATTCCAGTCATCAGGGC 357

Qy         130 CysSerMetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeu 149
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         358 TGCTCCATGTGGATGATACAGCACTGTGCCCGAGATGTTCTGGAGGCCCTCGCTTTTCTT 417

Qy         150 HisHisGluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGlu 169
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         418 CACCATGAGGGCTACGTCCATGCAGACCTCAAACCACGAAACATCCTGTGGAGTGCGGAG 477

Qy         170 AsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspVal 189
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         478 AATGAGTGTTTTAAGCTTATTGACTTTGGACTCAGCTTCAAAGAAGGCAATCAGGACGTA 537

Qy         190 LysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAla 209
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         538 AAGTATATTCAGACAGACGGGTATCGCGCTCCTGAAGCAGAGCTGCAGAACTGCTTGGCC 597

Qy         210 GlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGly 229
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         598 CAGGCCGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTCGGA 657

Qy         230 IleIleLeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGlu 249
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         658 ATCATTTTACTGGAAATGTTCTCAGGAATGAAGCTGAAACATACAGTCAGATCTCAGGAA 717

Qy         250 TrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAlaValValAsn 269
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         718 TGGAAGGCAAACAGTTCTGCTATTATTGATCATATATTTGCCAGTAAAGCAGTGGTGAAT 777

Qy         270 AlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspPro 289
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         778 GCCGCAATTCCAGCCTATCACCTCAGAGACCTTATCAAAAGCATGCTTCATGATGACCCG 837

Qy         290 SerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAla 309
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         838 GGCAGAAGAATCCCTGCTGAGATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCC 897

Qy         310 ProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeuAsnValLeu 329
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         898 CCTCATATTGAAGATCTGGTGATGCTTCCGACTCCAGTGCTCAGGCTCCTCAATGTGCTG 957

Qy         330 AspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAspValLysGlu 349
               |||||||||||||   ||||||:::||||||||||||||||||||||||||||||||||||
Db         958 GATGATGATTACCTTGAAAATGAAGATGAATATGAAGATGTTGTAGAAGATGTGAAAGAG 1017

Qy         350 GluCysGlnLysTyrGlyProValValSerLeuLeuValProLysGluAsnProGlyArg 369
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1018 GAGTGTCAGAAATATGGACCAGTGGTTTCTCTGCTTGTTCCAAAGGAAAATCCTGGCAGA 1077

Qy         370 GlyGlnValPheValGluTyrAlaAsnAlaGlyAspSerLysAlaAlaGlnLysLeuLeu 389
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1078 GGACAAGTCTTCGTTGAGTACGCGAACGCTGGTGATTCCAAAGCTGCTCAGAAGTTGCTG 1137

Qy         390 ThrGlyArgMetPheAspGlyLysPheValValAlaThrPheTyrProLeuSerAlaTyr 409
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1138 ACTGGGAGGATGTTTGACGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGCCTAC 1197

Qy         410 LysArgGlyTyrLeuTyrGlnThrLeuLeu 419
               |||||||||||||||||||||||||||||
Db        1198 AAGAGGGGATATCTTTATCAAACCTTGCTT 1227
```

```
RESULT 3
BM468107
LOCUS       BM468107                1079 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6432293 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5535653
            5', mRNA sequence.
ACCESSION   BM468107
VERSION     BM468107.1  GI:18517149
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12224  row: b  column: 06
            High quality sequence stop: 754.
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5535653"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

Alignment Scores:
Pred. No.:              5.61e-149       Length:         1079
Score:                  1521.50         Matches:        309
Percent Similarity:     92.86%          Conservative:   3
Best Local Similarity:  91.96%          Mismatches:     17
Query Match:            68.97%          Indels:         7
DB:                     4               Gaps:           3

US-10-798-532-2 (1-419) x BM468107 (1-1079)

Qy          90 ValThrLeuTyrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           2 GTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTG 61

Qy         110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGlnGly 129
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 TTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGT 121

Qy         130 CysSerMetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeu 149
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 TGTTCCATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTT 181

Qy         150 HisHisGluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGlu 169
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         182 CATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAG 241

Qy         170 AsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspVal 189
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 AATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTA 301

Qy         190 LysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAla 209
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCC 361

Qy        210 GlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGly 229
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGA 421

Qy        230 IleIleLeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGlu 249
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAA 481

Qy        250 TrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAlaValValAsn 269
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAAT 541

Qy        270 AlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspPro 289
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCA 601

Qy        290 SerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAla 309
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCC 661

Qy        310 ProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeuAsnValLeu 329
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTG 721

Qy        330 AspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAspValLysGlu 349
                ||||||||||||||||  ||||||||||||||||||||||||||||||:::  ||||||
Db        722 GATGATGATTATCTTGAGAATGAAGAGGAATATGAAGATGTTGTAGAAGAATGTAAAGAG 781

Qy        350 GluCysGlnLysTyrGlyProValValSerLeuLeuValProLysGluAsnProGlyArg 369
                ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db        782 GAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGNTCCAAAGGAAAATCCTGGCAGA 841

Qy        370 GlyGlnVal-PheValGluTyrAlaAsnAla-GlyAspSerLysAlaAlaGlnLys-Leu 388
                ||||||||| |||   ||||||||||||||    ||||||||||||||||||||| |||
Db        842 GGACAAGTCTTTTGNTGAGTATGCAAATGCTGNGTGATTCAAAAGCTGCGCAGAAATTTA 901

Qy        389 LeuThrGlyArg---MetPheAspGlyLysPheValValAla-ThrPheTyrProLeu-- 406
                |||||||||         |||||||||   ||||||  |||       |||
Db        902 CTGACTGGNAANGATGGTTTGATGGGGAAGTTTGNTGTNGCCTACAATCCTACCCGCCTG 961

Qy        407 -SerAlaTyrLysArgGly---TyrLeuTyrGlnThrLeu 418
                 ||||||   ::::::|||      ||||||     |||
Db        962 GAGTGCCTTACAGAAGGGGATTATTCTGTATTCAAACCTT 1001

RESULT 4
BY712344
LOCUS       BY712344               1015 bp    mRNA    linear   EST 17-DEC-2002
DEFINITION  BY712344 RIKEN full-length enriched, 10, 11 days embryo whole body
            Mus musculus cDNA clone 2810454O04 5', mRNA sequence.
ACCESSION   BY712344
VERSION     BY712344.1  GI:27123621
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1015)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
  MEDLINE   22354683
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
            Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
            Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
            Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
            Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
            Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .1015
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="2810454O04"
                     /tissue_type="whole body"
                     /dev_stage="10, 11 days embryo"
                     /clone_lib="RIKEN full-length enriched, 10, 11 days embryo
                     whole body"
ORIGIN

Alignment Scores:
Pred. No.:              4.99e-123       Length:         1015
Score:                  1274.50         Matches:        252
Percent Similarity:     89.51%          Conservative:   4
Best Local Similarity:  88.11%          Mismatches:     23
Query Match:            57.77%          Indels:         7
DB:                     6               Gaps:           4
```

```
US-10-798-532-2 (1-419) x BY712344 (1-1015)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 218

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 278

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              |||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 338

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 398

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 GANCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 458

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 518

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 578

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 638

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 698

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 758

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 818

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||   ||||||||||   ||||||||||||||||||||||||||||||
Db        819 TCAGCTGTTGATCTGTCGGGGCTCGGAATCACTTTACTGGAGATGTTCTCAGGAATGAAG 878

Qy        241 LeuLys---HisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
              ||||||   |||   :::           |||:::|||   ||||||
Db        879 CTGAAAATACATCAGATCTCAGGA------TGGGAGGCA---AGCAGTCTGCATCATTCG 929

Qy        260 HisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAsp 279
              |||||||||         ||||||   :::                 |||       |||
Db        930 CATATATTT---------GCAGTGAGCAGTGGTGGTNCGCAATTCAGCCATCACTCAGAC 980

Qy        280 LeuIleLysSerMetLeu 285
                 :::|||||||||
Db        981 GCTTATCAGAGCATGCTC 998

RESULT 5
BM451184
LOCUS       BM451184                 917 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6392719 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5495159
            5', mRNA sequence.
ACCESSION   BM451184
VERSION     BM451184.1  GI:18500224
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12121  row: j  column: 24
            High quality sequence stop: 718.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5495159"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
ORIGIN

Alignment Scores:
Pred. No.:                1.13e-119      Length:          917
Score:                    1242.00        Matches:         240
Percent Similarity:       99.59%         Conservative:    0
Best Local Similarity:    99.59%         Mismatches:      1
Query Match:              56.30%         Indels:          0
DB:                       4              Gaps:            0

US-10-798-532-2 (1-419) x BM451184 (1-917)

Qy        179 GlyLeuSerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArg 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 GGGCTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGG 81

Qy        199 AlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlu 218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAA 141

Qy        219 CysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGly 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 TGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGA 201

Qy        239 MetLysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIle 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 ATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATT 261

Qy        259 AspHisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArg 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGA 321

Qy        279 AspLeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAla 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCA 381

Qy        299 LeuCysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeu 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 TTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTT 441

Qy        319 ProThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGlu 338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        442 CCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAG 501
```

```
Qy        339 GluTyrGluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValVal 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 GAATATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTA 561

Qy        359 SerLeuLeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsn 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        562 TCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAAT 621

Qy        379 AlaGlyAspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPhe 398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        622 GCTGGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTT 681

Qy        399 ValValAlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeu 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        682 GTTGTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTG 741

Qy        419 Leu 419
              |||
Db        742 CTT 744
```

RESULT 6
CO771206
LOCUS CO771206 823 bp mRNA linear EST 04-AUG-2004
DEFINITION testis_EST02303 Testis cDNA Library Gallus gallus cDNA 3', mRNA sequence.
ACCESSION CO771206
VERSION CO771206.1 GI:50972473
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 823)
AUTHORS Shin,J.
TITLE Testis cDNA Library
JOURNAL Unpublished (2004)
COMMENT Contact: JiHye Shin
Division of Animal Genetic Engineering
School of Agricultural Biotechnology, Seoul National University
San 56-1,Sillim-dong, Gwanak-gu, Seoul, 151-742, Korea
Tel: +82-31-479-7557
Fax: +82-31-479-7550
Email: paecilojh@hanmail.net
Seq primer: T3.

FEATURES Location/Qualifiers
source 1. .823
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Korean Native chicken"
/db_xref="taxon:9031"
/sex="Male"
/tissue_type="Testis"
/dev_stage="26weeks"
/clone_lib="Testis cDNA Library"
/note="Vector: Uni-ZAP XR vector; Site_1: XhoI; Site_2: EcoRI; The libraries were synthesized using the ZAP(r)-cDNA synthesis method (Stratagene). cDNA was prepared, size-fractionated and inserted into the Uni-ZAP XR vector using XhoI linker-primer and EcoRI adaptor. After in vivo excision with E. coli strain SOLR, cDNA libraries from testis contain each insert size range of 0.5~3kb (n=20), and 0.5~3.5kb (n=20)."

ORIGIN

Alignment Scores:

Pred. No.:	3.61e-112	Length:	823
Score:	1170.00	Matches:	229
Percent Similarity:	84.72%	Conservative:	15
Best Local Similarity:	79.51%	Mismatches:	22
Query Match:	53.04%	Indels:	22
DB:	7	Gaps:	2

US-10-798-532-2 (1-419) x CO771206 (1-823)

```
Qy         47 GlySerProPro---------GlyAlaLeuLysGlnPheLeuProProGlyThrThrGly 63
              |||:::||||||          ||||||:::        |||||||||
Db          6 GGAGCTCCACCGCGGTGGCNGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAAT 65

Qy         64 AlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeuGluGlnLeu 83
              :::|||
Db         66 TCGGCACGAGGC---------------------------------------------- 77

Qy         84 GlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSerProAsn 103
                      ||||||||||||||||||||||||||||||||   |||:::|||   |||
Db         78 ---------CGCAACATCGTGACTCTGTACGGCGTGTTCACCAACCACTACTCGGCCAAC 128

Qy        104 ValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeu 123
              |||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||
Db        129 GGCCCGTCCCGCTGCCTGCTGCTGGAGCTGCTGGATATCAGCGTGTCTGAGCTGCTGCTG 188

Qy        124 TyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArgAspValLeu 143
              :::||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 CACTCCAGCAACCAGGGCTGCTCCATGTGGATGATCCAGCACTGCGCCCGCGACGTGCTG 248

Qy        144 GluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLysProArgAsn 163
              ||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||||
Db        249 GAAGCCCTGGCCTTCCTGCACCACAAAGGCTACGTGCACGCAGACCTGAAGCCACGCAAC 308

Qy        164 IleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLys 183
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        309 ATCCTGTGGAGCGCCGAGGAGGAGTGCTTTAAGCTCATTGACTTTGGACTTAGCTTCAAA 368

Qy        184 GluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGlu 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GAGGGGAATCAGGATGTGAAATATATTCAAACAGACGGGTATCGGGCTCCAGAGGCAGAA 428

Qy        204 LeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaVal 223
              ||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||
Db        429 CTGCAGAACTGCCTAGCACAGGCAGGGCTCCAGAGTGAGACGGAATGTACCTCTGCTGTG 488

Qy        224 AspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLysLeuLysHis 243
              |||||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||
Db        489 GATCTGTGGAGTCTGGGAATTGTTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACAT 548

Qy        244 ThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAla 263
              ||||||:::|||||||||||||||   ||||||||||||||||||||   |||||||||
Db        549 ACAGTCCAATCTCAGGAATGGAAGACAAACAGTTCTGCCATCATCGATCGCATTTTTGCC 608

Qy        264 SerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSer 283
              |||:::   |||||||||:::|||||||||||||||||||||||||||||||||||||||
Db        609 AGTGAAGGGGTGGTTAATTCAGCCATTCCAGCTTATCACCTCAGAGACCTTATTAAAAGC 668

Qy        284 MetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhe 303
              |||||||||   |||      :::|||      ||||||   ||||||||||||||||||||
Db        669 ATGCTTCATTGTGACCAAGGCAAGCGAGCCTCTGCTGAAAAGGCTTTATGCAGCCCATTC 728

Qy        304 PheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrProValLeu 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 TTCAGCATTCCCTTTGCTCCCCATATTGAAGATTTGGTGATGCTCCCCACGCCTGTGCTG 788

Qy        324 ArgLeuLeuAsnValLeuAspAsp 331
              ||||||||||||||||:::|||
Db        789 AGGCTGCTGAATGTTCTAAACGAT 812
```

RESULT 7
BU447831
LOCUS BU447831 779 bp mRNA linear EST 29-NOV-2002
DEFINITION 603767571F1 CSEQRBN13 Gallus gallus cDNA clone ChEST695n17 5', mRNA sequence.
ACCESSION BU447831
VERSION BU447831.1 GI:25937142
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 779)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .779
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Layer"
                     /db_xref="taxon:9031"
                     /clone="ChEST695n17"
                     /sex="Female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQRBN13"
                     /note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
                     EcoRI; Site_2: NotI; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with
                     EcoRI, size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer
                     reannealing hybridization was used."
ORIGIN

Alignment Scores:
Pred. No.:              2.31e-111       Length:         779
Score:                  1162.00         Matches:        224
Percent Similarity:     93.02%          Conservative:   16
Best Local Similarity:  86.82%          Mismatches:     18
Query Match:            52.67%          Indels:         1
DB:                     5               Gaps:           0

US-10-798-532-2 (1-419) x BU447831 (1-779)

Qy        158 AspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAsp 177
              ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db          6 GACCTGAAGCCACGCAACATCCTGTGGAGCGCGGAGGAGGAGTGCTTTAAGCTCATTGAC 65

Qy        178 PheGlyLeuSerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyr 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TTTGGACTTAGCTTCAAAGAGGGGAATCAGGATGTGAAATATATTCAAACAGACGGGTAT 125

Qy        198 ArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThr 217
              ||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        126 CGGGCTCCAGAGGCAGAACTGCAGAACTGCCTAGCACAGGCAGGGCTCCAGAGTGAGACG 185

Qy        218 GluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSer 237
              ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        186 GAATGTACCTCTGCTGTGGATCTGTGGAGTCTGGGAATTGTTTTACTGGAAATGTTCTCA 245

Qy        238 GlyMetLysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIle 257
              ||||||||||||||||||||||:::|||||||||||||||   ||||||||||||||||
Db        246 GGAATGAAACTGAAACATACAGTCCAATCTCAGGAATGGAAGACAAACAGTTCTGCCATC 305

Qy        258 IleAspHisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeu 277
              ||||||   ||||||||||||:::   |||||||||:::|||||||||||||||||||||
Db        306 ATCGATCGCATTTTTGCCAGTGAAGGGGTGGTTAATTCAGCCATTCCAGCTTATCACCTC 365

Qy        278 ArgAspLeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMet 297
              ||||||||||||||||||||||||||||||   |||      :::|||      ||||||
Db        366 AGAGACCTTATTAAAAGCATGCTTCATTGTGACCAAGGCAAGCGAGCCTCTGCTGAAAAG 425

Qy        298 AlaLeuCysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMet 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        426 GCTTTATGCAGCCCATTCTTCAGCATTCCCTTTGCTCCCCATATTGAAGATTCGGTGATG 485

Qy        318 LeuProThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGlu 337
              ||||||||||||||||||||||||||||||||||||||||   |||      |||   :::|||
Db        486 CTCCCCACGCCTGTGCTGAGGCTGCTGAATGTTCTAAGCGATGCTTCTCTGCAGAGTGAA 545

Qy        338 GluGluTyrGluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProVal 357
              ||||||||||||||||:::::|||||:::::|||||||||||||||||||||||||||||
Db        546 GAAGAATACGAAGATATCCTGGAAGACATAAGGGAGGAGTGTCAGAAATATGGACCGGTG 605

Qy        358 ValSerLeuLeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAla 377
              ||||||||||||:::|||||||||||||||||:::|||||||||||   |||||||||
Db        606 GTTTCCTTGCTTATTCCGAAGGAAAATCCTGGTAAAGGCCAAGTCTTTGT-GAATATGCA 664

Qy        378 AsnAlaGlyAspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLys 397
              ||||||||||||||||||||||||||:::|||||||||:::::|||||||||||||||||
Db        665 AATGCTGGTGATTCCAAAGCTGCCCAAAAAATGCTGACTGGAAAGATTTTTGATGGCAAG 724

Qy        398 PheValValAlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyr 415
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        725 TCTGTCGTGGCTACGTNTTACCCACTGAGTGCCTATAAGAGAGGATATCTGTAC 778

RESULT 8
CO773091
LOCUS       CO773091                896 bp    mRNA    linear   EST 04-AUG-2004
DEFINITION  testis_EST04188 Testis cDNA Library Gallus gallus cDNA 3', mRNA
            sequence.
ACCESSION   CO773091
VERSION     CO773091.1  GI:50974358
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 896)
  AUTHORS   Shin,J.
  TITLE     Testis cDNA Library
  JOURNAL   Unpublished (2004)
COMMENT     Contact: JiHye Shin
            Division of Animal Genetic Engineering
            School of Agricultural Biotechnology, Seoul National University
            San 56-1,Sillim-dong, Gwanak-gu, Seoul, 151-742, Korea
            Tel: +82-31-479-7557
            Fax: +82-31-479-7550
            Email: paecilojh@hanmail.net
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .896
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Korean Native chicken"
                     /db_xref="taxon:9031"
                     /sex="Male"
                     /tissue_type="Testis"
                     /dev_stage="26weeks"
                     /clone_lib="Testis cDNA Library"
                     /note="Vector: Uni-ZAP XR vector; Site_1: XhoI; Site_2:
                     EcoRI; The libraries were synthesized using the
                     ZAP(r)-cDNA synthesis method (Stratagene). cDNA was
                     prepared, size-fractionated and inserted into the Uni-ZAP
```

XR vector using XhoI linker-primer and EcoRI adaptor. After in vivo excision with E. coli strain SOLR, cDNA libraries from testis contain each insert size range of 0.5~3kb (n=20), and 0.5~3.5kb (n=20)."

ORIGIN

Alignment Scores:

Pred. No.:	5.89e-111	Length:	896
Score:	1159.00	Matches:	240
Percent Similarity:	81.90%	Conservative:	18
Best Local Similarity:	76.19%	Mismatches:	30
Query Match:	52.54%	Indels:	27
DB:	7	Gaps:	2

US-10-798-532-2 (1-419) x CO773091 (1-896)

```
Qy         46 ProGlySer-ProPro---------GlyAlaLeuLysGlnPheLeuProProGlyThrTh 62
              ||||||::: ||||||          ||||||:::       |||||||||
Db         11 CCTGGNACCTCCACCGCGGTGGCGNGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAG 70

Qy         62 rGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeuGluGl 82
                  :::|||
Db         71 GAATTCGGCACGAGGC---------------------------------------------- 86

Qy         82 nLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSerPr 102
                        ||||||||||||||||||||||||||||||||   |||:::|||
Db         87 -------------CGCAACATCGTGACTCTGTACGGCGTGTTCACCAACCACTACTCGGC 133

Qy        102 oAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGluLeuLe 122
               |||   ||||||||||||||||||||||||||||||:::|||||||||||||||||
Db        134 CAACGGCCCGTCCCGCTGCCTGCTGCTGGAGCTGCTGGATATCAGCGTGTCTGAGCTGCT 193

Qy        122 uLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArgAspVa 142
              ||||:::||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGCACTCCAGCAACCAGGGCTGCTCCATGTGGATGATCCAGCACTGCGCCCGCGACGT 253

Qy        142 lLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLysProAr 162
              ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db        254 GCTGGAAGCCCTGGCCTTCCTGCACCACAAAGGCTACGTGCACGCAGACCTGAAGCCACG 313

Qy        162 gAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeuSerPh 182
              ||||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db        314 CAACATCCTGTGGAGCGCCGAGGAGGAGTGCTTTAAGCTCATTGACTTTGGACTTAGCTT 373

Qy        182 eLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaProGluAl 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 CAAAGAGGGGAATCAGGATGTGAAATATATTCAAACAGACGGGTATCGGGCTCCAGAGGC 433

Qy        202 aGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAl 222
              ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||
Db        434 AGAACTGCAGAACTGCCTAGCACAGGCAGGGCTCCAGAGTGAGACGGAATGTACCTCTGC 493

Qy        222 aValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLysLeuLy 242
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        494 TGTGGATCTGTGGAGTCTGGGAATTGTTTTACTGGAAATGTTCTCAGGAATGAAACTGAA 553

Qy        242 sHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePh 262
              ||||||||||:::|||||||||||||||   ||||||||||||||||||||   |||||
Db        554 ACATACAGTCCAATCTCAGGAATGGAAGACAAACAGTTCTGCCATCATCGATCGCATTTT 613

Qy        262 eAlaSer-LysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIleL 282
              ||||||| |||   |||||||||:::|||||||||||||||||||||||||||||||||||
Db        614 TGCCAGTGAAGGGGGTGGTTAATTCAGCCATTCCAGCTTATCACCTCAGAGACCTTATTA 673

Qy        282 ysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerP 302
              ||||||||||||||   |||      :::|||      ||||||   ||||||||||||||
Db        674 AAAGCATGCTTCATTGTGACCAAGGCAAGCGAGCCTCTGCTGAAAAGGCTTTATGCAGCC 733

Qy        302 roPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrProV 322
              ||||||||||||||||||||||||||||::::::||||||||||||||||||||||||||
Db        734 CATTCTTCAGCATTCCCTTTGCTCCCCATATGAAAGATTTGGTGATGCTCCCCACGCCTG 793

Qy        322 alLeuArgLeu-LeuAsnValLeuAspAspAspTyrLeuGlyAsn-GluGluGluTyrGl 341
```

```
              |||||||||||| ||||||||||||   |||      |||   ::: ||||||||||||||
Db        794 TGCTGAGGCTGNCTGAATGTTCTAAGCGATGCTTCTCTGCAGAGTGGAAGAAGAATACGA 853

Qy        341 uAspValValGlu-AspValLysGluGluCysGln 352
                  |:::   :::|||          ||||||||||
Db        854 AGAATCCCTGGAGACATAGGGNAGAGGAGTGTCAG 888
```

RESULT 9
CN232453

LOCUS	CN232453 781 bp mRNA linear EST 09-APR-2004
DEFINITION	WLB074D11.ab1 WLtestis Gallus gallus cDNA 5', mRNA sequence.
ACCESSION	CN232453
VERSION	CN232453.1 GI:46336197
KEYWORDS	EST.
SOURCE	Gallus gallus (chicken)
ORGANISM	Gallus gallus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE	1 (bases 1 to 781)
AUTHORS	Savolainen,P., Fitzsimmons,C.J., Arvestad,L., Andersson,L. and Lundeberg,J.
TITLE	EST analysis of brain and testis cDNA libraries from White leghorn and Red Jungle Fowl
JOURNAL	Unpublished (2004)
COMMENT	Contact: Peter Savolainen Department of Biotechnology Royal Institute of Technology, KTH SE-106 91 Stockholm, SWEDEN Tel: +46 (0)8 5537 8481 Fax: +46 (0)8 5537 8335 Email: Peter.Savolainen@biotech.kth.se Seq primer: M13 reverse primer.

FEATURES Location/Qualifiers

```
     source          1. .781
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="White Leghorn"
                     /db_xref="taxon:9031"
                     /sex="male"
                     /lab_host="ElectroMAX DH10B (Invitrogen)"
                     /clone_lib="WLtestis"
                     /note="Organ: testis; Vector: pSPORT-1; Site_1: Hind III;
                     Site_2: EcoRI; The cDNA libraries were created with the
                     Superscript Plasmid System (Invitrogen)."
```

ORIGIN

Alignment Scores:

Pred. No.:	2.61e-108	Length:	781
Score:	1133.00	Matches:	225
Percent Similarity:	93.10%	Conservative:	18
Best Local Similarity:	86.21%	Mismatches:	18
Query Match:	51.36%	Indels:	3
DB:	7	Gaps:	0

US-10-798-532-2 (1-419) x CN232453 (1-781)

```
Qy        109 LeuLeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGln 128
              ||||||||||||||||||||:::|||||||||||||||||||||:::|||||||:::|||
Db          2 CTGCTGCTGGAGCTGCTGGATATCAGCGTGTCTGAGCTGCTGCTGCACTCCAGCAACCAG 61

Qy        129 GlyCysSerMetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPhe 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GGCTGCTCCATGTGGATGATCCAGCACTGCGCCCGCGACGTCCTGGAAGCCCTGGCCTTC 121

Qy        149 LeuHisHisGluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAla 168
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CTGCACCACAAAGGCTACGTGCACGCAGACCTGAAGCCACGCAACATCCTGTGGAGCGCG 181

Qy        169 GluAsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAsp 188
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GAGGAGGAGTGCTTTAAGCTCATTGACTTTGGACTTAGCTTCAAAGAGGGGAATCAGGAT 241
```

```
Qy        189 ValLysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeu 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GTGAAATATATTCAAACAGACGGGTATCGGGCTCCAGAGGCAGAACTGCAGAACTGCCTA 301

Qy        209 AlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeu 228
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        302 GCACAGGCAGGGCTCCAGAGTGAGACGGAATGTACCTCTGCTGTGGATCTGTGGAGTCTG 361

Qy        229 GlyIleIleLeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGln 248
              ||||||:::||||||||||||||||||||||||||||||||||||:::||||||||||||||
Db        362 GGAATTGTTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCCAATCTCAG 421

Qy        249 GluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAlaValVal 268
              |||||||||   ||||||||||||||||||||   ||||||||||||:::   ||||||
Db        422 GAATGGAAGACAAACAGTTCTGCCATCATCGATCGCATTTTTGCCAGTGAAGGGGTGGTT 481

Qy        269 AsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAsp 288
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        482 AATTCAGCCATTCCAGCTTATCACCTCAGAGACCTTATTAAAAGCATGCTTCATTGTGAC 541

Qy        289 ProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPhe 308
                    :::|||      ||||||  ||||||||||||||||||||||||||||||||||||
Db        542 CAAGGCAAACGAGCCTCTGCTGAAAAGGCTTTATGCAGCCCATTCTTCAGCATTCCCTTT 601

Qy        309 AlaProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeuAsnVal 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GCTCCCCATATTGAAGATTTGGTGATGCTCCCCACGCCTGTGCTGAGGCTGCTGAATGTT 661

Qy        329 LeuAspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAspValLys 348
              |||   |||      |||   :::|||   ||||||||||:::::|||   ::::::
Db        662 CTAAGCGATGCTTCTCTGCAGAGTGAAGA-GAATACGAAGATATCCTGGAAGA-ATAAGG 719

Qy        349 GluGluCysGlnLysTyrGlyProValValSerLeuLeuValProLysGluAsnProGly 368
              ||||||||||||||||||||||||   |||||||||:::|||:::   |||||||||
Db        720 GAGGAGTGTCAGAAATATGGACCGGTGGGTTCCTTGCTTATTCCGGAGGA-AATCCTGGT 778

Qy        369 Arg 369
              :::
Db        779 AAA 781

RESULT 10
BU541306
LOCUS       BU541306                 932 bp    mRNA    linear   EST 13-SEP-2002
DEFINITION  AGENCOURT_10327419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6572970
            5', mRNA sequence.
ACCESSION   BU541306
VERSION     BU541306.1  GI:22851747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 932)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2765  row: o  column: 18
            High quality sequence stop: 631.
FEATURES             Location/Qualifiers
     source          1. .932
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6572970"
                     /tissue_type="carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_40"
                     /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed  by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California,  Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life  Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN
```

Alignment Scores:

Pred. No.:	8.94e-106	Length:	932
Score:	1110.00	Matches:	228
Percent Similarity:	87.22%	Conservative:	4
Best Local Similarity:	85.71%	Mismatches:	18
Query Match:	50.32%	Indels:	16
DB:	5	Gaps:	2

US-10-798-532-2 (1-419) x BU541306 (1-932)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 208

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 268

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 328

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 388

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 448

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 508

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        509 TTGCTCTTATATTCCAGTCACCAGCGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 568

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeu-Ly 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        569 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCCAA 628

Qy        160 sProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLe 180
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        629 ACCACGTAACATATTGTGGAGTGCAAAGAATGAATGTTTTAAACTCATTGACTTTGGACT 688

Qy        180 u-SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaP 200
              | ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        689 TTAGCTTCAAAGAAAGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTC 748

Qy        200 roGlu-AlaGluLeuGlnAsnCysLeuAlaGlnAla-GlyLeuGlnSerAspThrGluCy 219
              |||||    ||||||||||||||||||||    ||| ||||||||||||||||:::||
Db        749 CAGAAAACAGAATTGCAAAATTGCTTGGCCCANGCTTGGCCTGCAAAGTGATACAAAATG 808

Qy        219 sThrSer-AlaValAsp-LeuTrpSer------LeuGlyIleIleLeuLeuGluMetPhe 236
              ||||||| ||||||||| |||||||||                  |||
Db        809 TACCTCCAGCTGTTGATCTTGTGGAGCCCTAAGAAATCATTTTACCTGGAAAATGGTCTC 868

Qy        237 SerGlyMetLysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAla 256
                 ||||||                        :::   ||||||         |||
```

```
Db        869 CAGGGAATG----------------------GAAAACTGGAAAACCTACCGGTCCAGA 904

Qy        257 IleIle 258
                |||:::
Db        905 ATCCTC 910

RESULT 11
BU385761
LOCUS       BU385761                717 bp    mRNA    linear   EST 28-NOV-2002
DEFINITION  603582650F1 CSEQCHN75 Gallus gallus cDNA clone ChEST533d13 5', mRNA
            sequence.
ACCESSION   BU385761
VERSION     BU385761.1  GI:25893762
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 717)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .717
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="White Leghorn, Hisex"
                     /db_xref="taxon:9031"
                     /clone="ChEST533d13"
                     /dev_stage="36"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN75"
                     /note="Organ: trunks; Vector: pBluescript II KS(+);
                     Site_1: EcoRI; Site_2: Not1; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with
                     EcoRI, size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer
                     reannealing hybridization was used."
ORIGIN

Alignment Scores:
Pred. No.:              1.59e-105       Length:         717
Score:                  1106.00         Matches:        210
Percent Similarity:     93.72%          Conservative:   14
Best Local Similarity:  87.87%          Mismatches:     14
Query Match:            50.14%          Indels:         1
DB:                     5               Gaps:           0

US-10-798-532-2 (1-419) x BU385761 (1-717)

Qy        134 MetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeuHisHisGluGly 153
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db          3 ATGATCCAGCACTGCGCCCGCGACGTCCTGGAAGCCCTGGCCTTCCTGCACCACAAAGGC 62

Qy        154 TyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGluCysPhe 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db         63 TACGTGCACGCAGACCTGAAGCCACGCAACATCCTGTGGAGCGCGGAGGAGGAGTGCTTT 122

Qy        174 LysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspValLysTyrIleGln 193
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 AAGCTCATTGACTTTGGACTTAGCTTCAAAGAGGGGAATCAGGATGTGAAATATATTCAA 182

Qy        194 ThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeu 213
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ACAGACGGGTATCGGGCTCCAGAGGCAGAACTGCAGAACTGCCTAGCACAGGCAGGGCTC 242

Qy        214 GlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeu 233
              ||||||:::|||||||||||||||||||||||||||||||||||||||:::||||||
Db        243 CAGAGTGAGACGGAATGTACCTCTGCTGTGGATCTGTGGAGTCTGGGAATTGTTTTACTG 302

Qy        234 GluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsn 253
              |||||||||||||||||||||||||||||||:::|||||||||||||||   |||
Db        303 GAAATGTTCTCAGGAATGAAACTGAAACATACAGTCCAATCTCAGGAATGGAAGACAAAC 362

Qy        254 SerSerAlaIleIleAspHisIlePheAlaSerLysAlaValValAsnAlaAlaIlePro 273
              |||||||||||||||||   |||||||||||:::   |||||||||:::|||||||||
Db        363 AGTTCTGCCATCATCGATCGCATTTTTGCCAGTGAAGGGGTGGTTAATTCAGCCATTCCA 422

Qy        274 AlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspProSerArgArgIle 293
              |||||||||||||||||||||||||||||||||   |||      :::|||
Db        423 GCTTATCACCTCAGAGACCTTATTAAAAGCATGCTTCATTGTGACCAAGGCAAACGAGCC 482

Qy        294 ProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaProHisIleGlu 313
                  ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TCTGCTGAAAAGGCTTTATGCAGCCCATTCTTCAGCATTCCCTTTGCTCCCCATATTGAA 542

Qy        314 AspLeuValMetLeuProThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyr 333
              |||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        543 GATTTGGTGATGCTCCCCACGCCTGTGCTGAGGCTGCTGAATGTTCTAAGCGATGCTTCT 602

Qy        334 LeuGlyAsnGluGluGluTyrGluAspValValGluAspValLysGluGluCysGlnLys 353
              |||   :::|||||||||||||||||:::::|||||:::::|||||||||||||||||
Db        603 CTGCAGAGTGAAGAAGAATACGAAGATATCCTGGAAGACATAAGGGAGGAGTGTCAGAAA 662

Qy        354 TyrGlyProValValSerLeuLeuValProLysGlu-AsnProGlyArgGlyGln 371
              |||||||||||||||||||||:::|||||||||| ||||||||:::|||||||
Db        663 TATGGACCGGTGGTTTCCTTGCTTATTCCGAAGGAACAATCCTGGTAAAGGCCAA 717

RESULT 12
BE895119
LOCUS       BE895119                899 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601436041F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920939 5',
            mRNA sequence.
ACCESSION   BE895119
VERSION     BE895119.1  GI:10358193
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 899)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9753  row: b  column: 12
            High quality sequence stop: 626.
FEATURES             Location/Qualifiers
```

```
     source          1. .899
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3920939"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

Alignment Scores:
Pred. No.:              8.45e-104       Length:         899
Score:                  1091.00         Matches:        210
Percent Similarity:     99.53%          Conservative:   0
Best Local Similarity:  99.53%          Mismatches:     0
Query Match:            49.46%          Indels:         1
DB:                     2               Gaps:           0

US-10-798-532-2 (1-419) x BE895119 (1-899)

Qy        113 LeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGlnGlyCysSerMet 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATG 60

Qy        133 TrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeuHisHisGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAG 120

Qy        153 GlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGluCys 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGT 180

Qy        173 PheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspValLysTyrIle 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATT 240

Qy        193 GlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGly 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGC 300

Qy        213 LeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeu 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTA 360

Qy        233 LeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGluTrpLysAla 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCA 420

Qy        253 AsnSerSerAlaIleIleAspHisIlePheAlaSerLysAlaValValAsnAlaAlaIle 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATT 480

Qy        273 ProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspProSerArgArg 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGA 540

Qy        293 IleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaProHisIle 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATT 600

Qy        313 GluAspLeuValMetLeu-ProThrProVal 322
              |||||||||||||||| ||||||||||||
Db        601 GAAGATCTGGTCATGCTTTCCCACTCCAGTT 631

RESULT 13
BI562420
LOCUS       BI562420                 772 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603256316F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
            mRNA sequence.
ACCESSION   BI562420
VERSION     BI562420.1  GI:15449747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 772)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11755  row: p  column: 03
            High quality sequence stop: 727.
FEATURES             Location/Qualifiers
     source          1. .772
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5298674"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTVN-3', size-selected for  average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation).  Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.23e-102       Length:         772
Score:                  1079.00         Matches:        212
Percent Similarity:     96.82%          Conservative:   1
Best Local Similarity:  96.36%          Mismatches:     5
Query Match:            48.91%          Indels:         3
DB:                     4               Gaps:           1

US-10-798-532-2 (1-419) x BI562420 (1-772)

Qy         92 LeuTyrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeuLeuLeu 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 TTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTGTTGCTT 167

Qy        112 GluLeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGlnGlyCysSer 131
              |||||||||||||||||||||||||||   ||||||||||||||||   |||||||||
Db        168 GAACTCCTGGATGTCAGTGTTTCGGAANTGCTCTTATATTCCAGTCACCANGGTTGTTCC 227

Qy        132 MetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeuHisHis 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 ATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCAT 287

Qy        152 GluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGlu 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAA 347

Qy        172 CysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspValLysTyr 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 TGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 407
```

```
Qy        192 IleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAla 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 ATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCT 467

Qy        212 GlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIle 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 527

Qy        232 LeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGluTrpLys 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 587

Qy        252 AlaAsnSerSerAlaIleIleAspHisIle-PheAlaSerLysAlaValValAsnAlaAl 271
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        588 GCAAACAGTTCTGCTATTATTGATCACATTATTTGCCAGTAAAGCAGTGGTGAATGCCGC 647

Qy        271 aIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspProSerAr 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 AATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAG 707

Qy        291 gArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaPro 310
              |||||||||||||||||||||||||||  ||||||   :::      |||||||||
Db        708 AAGAATTCCTGCTGAAATGGCATTGTGCAG-CCATTCTTAGCA---TCCTTTGCCCCT 761

RESULT 14
BU205357
LOCUS       BU205357                794 bp    mRNA    linear   EST 25-NOV-2002
DEFINITION  603104569F1 CSEQCHN03 Gallus gallus cDNA clone ChEST41j18 5', mRNA
            sequence.
ACCESSION   BU205357
VERSION     BU205357.1  GI:25370848
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 794)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .794
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="White Leghorn, Hisex"
                     /db_xref="taxon:9031"
                     /clone="ChEST41j18"
                     /tissue_type="whole embryo"
                     /dev_stage="20-21"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN03"
                     /note="Organ: whole embryo; Vector: pBluescript II KS(+);
                     Site_1: EcoRI; Site_2: NotI; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with
                     EcoRI, size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer
                     reannealing hybridization was used."
ORIGIN

Alignment Scores:
Pred. No.:                3.38e-102       Length:          794
Score:                    1075.00         Matches:         221
Percent Similarity:       89.10%          Conservative:    16
Best Local Similarity:    83.08%          Mismatches:      27
Query Match:              48.73%          Indels:          4
DB:                       5               Gaps:            0

US-10-798-532-2 (1-419) x BU205357 (1-794)

Qy         82 GlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSer 101
              ||||||:::||||||||||||||||||||||||||||||||||||||||   |||:::|||
Db          1 CAGCTCCGCGGGCACCGCAACATCGTGACTCTGTACGGCGTGTTCACCAACCACTACTCG 60

Qy        102 ProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGluLeu 121
                 |||  ||||||||||||||||||||||||||||||:::|||||||||||||||||
Db         61 GCCAACGGCCCGTCCCGCTGCCTGCTGCTGGAGCTGCTGGATATCAGCGTGTCTGAGCTG 120

Qy        122 LeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArgAsp 141
              ||||||:::|||||||:::|||   ||||||||||||||||||||||||||||||||||||
Db        121 CTGCTGCACTCCAGCAACCAGGTCTGCTCCATGTGGATGATCCAGCACTGCGCCCGCGAC 180

Qy        142 ValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLysPro 161
              ||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        181 GTCCTGGAAGCCCTGGCCTTCCTGCACCACAAAGGCTACGTGCACGCAGACCTGAAGCCA 240

Qy        162 ArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeuSer 181
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        241 CGCAACATCCTGTGGAGCGCGGAGGAGGAGTGCTTTAAGCTCATTGACTTTGGACTTAGC 300

Qy        182 Phe-LysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaProGl 201
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCAAAAGAGGGGAATCAGGATGTGAAATATATTCAAACAGACGGGTATCGGGCTCCAGA 360

Qy        201 uAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSe 221
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db        361 GGCAGAACTGCAGAACTGCCTAGCACAGGCAGGGCTCCAGAGTGAGACGGAATGTACCTC 420

Qy        221 rAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLysLe 241
              ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db        421 TGCTGTGGATCTGTGGAGTCTGGGAATTGTTTTACTGGAAATGTTCTCAGGAATGAAACT 480

Qy        241 uLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIl 261
              |||||||||||||:::|||||||||||||||   ||||||||||||||||||||   ||
Db        481 GAAACATACAGTCCAATCTCAGGAATGGAAGACAAACAGTTCTGCCATCATCGATCGCAT 540

Qy        261 ePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIl 281
              ||||||||||:::   |||||||||  ||||||||||||||||||||||||||||||||||
Db        541 TTTGCCAGTGAAGGGGTGGTTAATCCAGCCATTCCAGCTTATCACCTCAGAGACCTTAT 600

Qy        281 eLysSerMetLeuHisAspAsp-ProSerArgArgIleProAlaGluMetAlaLeuCysS 301
              ||||||||||||||||   |||    ::::::|||      ||||||   |||   ||||
Db        601 TAAAAGCATGCTTCATTGTGACCAAGGCAAAACGAGCCTCTGCTGAAAAGGCTTAATGCA 660

Qy        301 erProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrP 321
              |||||||||||||||||||||||   |||   ||||||||||||||||||||||||||   |
Db        661 GCCCATTCTTCAGCATTCCCTTGGCTCC-CATATTGAAGATTTGGTGATGCTCCCAAG-C 718

Qy        321 roValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyrG 341
              ||||||||||||||||||||||||||||||   :::          :::||||
Db        719 CTGTGCTGAGGCTGCTGAATGTTCTAAGGCGAAGGTTCTCTGCAGAGTGAAGAAAATACG 778

Qy        341 luAspValValGlu 345
              ||::::::   |||
Db        779 AGAATATCTGGGAA 792

RESULT 15
```

```
CF593363
LOCUS       CF593363                834 bp    mRNA    linear   EST 26-SEP-2003
DEFINITION  AGENCOURT_15622279 NIH_MGC_147 Homo sapiens cDNA clone
            IMAGE:30531268 5', mRNA sequence.
ACCESSION   CF593363
VERSION     CF593363.1  GI:36346826
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Stefan Hansson
             cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
            and advice from Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM621  row: m  column: 05
            High quality sequence stop: 660.
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30531268"
                     /tissue_type="Human Placenta"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_147"
                     /note="Organ: placenta; Vector: pBluescriptR; Site_1:
                     alI-XhoI; Site_2: BamH; Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTTVN-3', size-selected  for average
                     insert size 2.3 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using  the Cap-trapper method (Carninci, in
                     preparation). Library constructed  by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: This is
                     a NIH_MGC library."
ORIGIN

Alignment Scores:
Pred. No.:              5.24e-102       Length:         834
Score:                  1073.50         Matches:        214
Percent Similarity:     95.13%          Conservative:   1
Best Local Similarity:  94.69%          Mismatches:     5
Query Match:            48.66%          Indels:         6
DB:                     7               Gaps:           1

US-10-798-532-2 (1-419) x CF593363 (1-834)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 214

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 274

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 334

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 394

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 454

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 514

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        515 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 574

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        575 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 634

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGly-Le 180
              |||||||||||||||||||||||||||||||||||   ||||||||||||||| ||
Db        635 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTANACTCATTGACTTTGGACTT 694

Qy        180 uSerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArg-AlaP 200
              |   |||||||||  ||||||||||||||||||||||||||||||||||||| ||||
Db        695 AGGCTTCAAAGAAAGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGGCTC 754

Qy        200 roGluAlaGluLeuGln-AsnCysLeuAla---GlnAlaGly-LeuGlnSerAspThrGl 218
              |||||||||||||||| |||||||||      :::   ||| ||||||||||||||||
Db        755 CAGAAGCAGAATTGCAAAAATTGCTTGGGCCCCAGGCTTGGCCCTGCAGAGTGATACAGA 814

Qy        218 u-CysThr 220
              | ||||||
Db        815 AATGTACC 822


Search completed: November 29, 2004, 13:27:43
Job time : 3172.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 26 Seconds
(without alignments)
1550.570 Million cell updates/sec

Title: US-10-798-532-2
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAFG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1052	47.7	200	2	I48615	gene KIS protein -
2	248.5	11.3	948	2	T24445	hypothetical prote
3	243.5	11.0	720	2	A96807	hypothetical prote
4	243.5	11.0	826	2	JC5153	mitogen-activated
5	241.5	10.9	387	2	T37758	protein kinase skp
6	240	10.9	388	2	S48879	protein kinase SMK
7	235.5	10.7	367	2	S68680	stress-activated p
8	234.5	10.6	354	2	T45138	protein kinase skp
9	234.5	10.6	433	2	S37790	probable serine/th
10	234	10.6	435	2	S64950	protein kinase HOG
11	233	10.6	298	1	A44878	protein kinase (EC
12	233	10.6	298	2	A41227	protein kinase (EC
13	233	10.6	369	2	A56492	protein kinase ERK
14	230	10.4	373	2	T13024	probable protein k
15	230	10.4	376	2	S40471	mitogen-activated
16	230	10.4	427	2	JC5693	stress-activated p
17	228	10.3	423	2	S43968	p54-alpha-2 stress
18	227.5	10.3	469	1	S17656	Ca2+/calmodulin-de
19	227	10.3	326	2	S23387	protein kinase (EC
20	227	10.3	376	2	S40470	mitogen-activated
21	227	10.3	423	2	S43967	p54-alpha stress-a
22	227	10.3	424	2	S71102	protein kinase JNK
23	227	10.3	529	1	S71774	calcium-dependent
24	226.5	10.3	367	2	JC5252	mitogen-activated
25	226	10.2	346	1	I48157	protein kinase (EC
26	226	10.2	406	1	KIHUCT	phosphorylase kina
27	225.5	10.2	301	1	S19209	protein kinase (EC
28	225.5	10.2	371	2	T14915	mitogen-activated
29	225.5	10.2	416	2	A48249	pre-mRNA splicing
30	224.5	10.2	301	1	A48041	protein kinase (EC
31	224.5	10.2	678	2	T43539	spindle checkpoint
32	224	10.2	358	2	S23383	protein kinase (EC
33	223.5	10.1	335	2	T23050	hypothetical prote
34	223.5	10.1	342	2	T21098	protein kinase (EC
35	223.5	10.1	426	2	S43969	p54-beta stress-ac
36	223.5	10.1	464	2	S71104	protein kinase JNK
37	223.5	10.1	602	2	S60052	calcium-dependent
38	223	10.1	427	2	JC5694	stress-activated p
39	223	10.1	575	2	JC7794	lammer kinase homo
40	223	10.1	690	2	T38052	probable protein k
41	222.5	10.1	393	2	S51321	mitogen-activated
42	222.5	10.1	424	2	A55480	c-Jun amino-termin
43	222.5	10.1	473	1	A53036	Ca2+/calmodulin-de
44	222	10.1	346	1	I78840	protein kinase (EC
45	222	10.1	371	2	S60121	mitogen-activated

ALIGNMENTS

RESULT 1
I48615
gene KIS protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I48615
R;Maucuer, A.; Camonis, J.H.; Sobel, A.
Proc. Natl. Acad. Sci. U.S.A. 92, 3100-3104, 1995
A;Title: Stathmin interaction with a putative kinase and coiled-coil-forming protein dom
A;Reference number: I48282; MUID:95241452; PMID:7724523
A;Accession: I48615
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-200 <RES>
A;Cross-references: UNIPROT:P97343; EMBL:X82320; NID:g791078; PIDN:CAA57763.1; PID:g7910
C;Genetics:
A;Gene: KIS

Query Match 47.7%; Score 1052; DB 2; Length 200;
Best Local Similarity 99.0%; Pred. No. 1.3e-53;
Matches 198; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      122 LLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLS 181
            ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db        1 LLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAEDECFKLIDFGLS 60

Qy      182 FKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL 241
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 FKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL 120

Qy      242 KHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALCS 301
            |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      121 KHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALCS 180

Qy      302 PFFSIPFAPHIEDLVMLPTP 321
            ||||||||||||||||||||
Db      181 PFFSIPFAPHIEDLVMLPTP 200
```

RESULT 2
T24445
hypothetical protein T04C10.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24445
R;Burton, J.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19891
A;Accession: T24445
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-948 <WIL>

```
A;Cross-references: UNIPROT:Q8WQL7; EMBL:Z69885; PIDN:CAA93756.1; GSPDB:GN00028; CESP:TC
A;Experimental source: clone T04C10
C;Genetics:
A;Gene: CESP:T04C10.1
A;Map position: X
A;Introns: 64/3; 181/3; 214/3; 238/1; 272/3; 331/2; 391/2; 457/1; 726/3; 830/2; 923/3

  Query Match            11.3%;  Score 248.5;  DB 2;  Length 948;
  Best Local Similarity   25.7%;  Pred. No. 6.2e-07;
  Matches  104;  Conservative   57;  Mismatches  113;  Indels  131;  Gaps   17;

Qy         88 NIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALA 147
              |||||  | | :| :      ||: |||  :: :||  :| :| |: : :  |: : : |
Db        455 NIVTLKGHF-VHRA----HLCLVFELLSYNLYDLLKNTSFRGVSLNLARKFAQQLGKTLL 509

Qy        148 FLHHE--GYVHADLKPRNILW-SAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAEL 204
              ||       :| |||| |:|  :|:    ::|||| | : |::  :|||:  ||:||  |
Db        510 FLSSPELSIIHCDLKPENVLLVNAKRSQIRVIDFGSSCQTGHRIYQYIQSRFYRSPEVLL 569

Qy        205 QNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSG-------------MKL---------- 241
                    |:  ||:    :|:|||| ||:|| :|              ||:
Db        570 -------GIAYDTK----IDMWSLGCILVEMHTGEPLFAGSSEVDQMMKIVEVLGMPPKE 618

Qy        242 ----------------------------KHTVRSQEWKANSSAIIDHIFASKAVVNAAI 272
                                         :||     :||  :  :  |      | :
Db        619 MLDIGPKTHKYFDKTEDGIYYCKKTRDGYRHT-----YKAPGARKLHEILG----VTSGG 669

Qy        273 PA-------------YHLRDLIKSMLHDDPSRRIPAEMALCSPFFSI--PFAPHIEDLV 316
              |                :|||| ||  || :||       ||:: :  ||    |: |
Db        670 PGGRRLGEPGHSVEDYSKFKDLIKRMLQFDPKQRI-------SPYYVVRHPFLKQKEERV 722

Qy        317 MLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKY--GPVVSLLVPKENPGRGQVFV 374
                 ||                        :::::  | |   |   :    |:|  | |:|
Db        723 PSQPPVSH-------------------SNLQQQQQLYIQQPSSQMSQVMESPSVGSVYV 762

Qy        375 E----YANAGDSKAAQKLLTGRMFDGKF--VVATFYPLSAYKRGY 413
              |    |  | | |    :|    :|    | |      ||::: |
Db        763 EDNGMYRQAPGSSANPISVTSSFDEGDAMEVDAGRRRFSAHQQNY 807


RESULT 3
A96807
hypothetical protein T32E8.5 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A96807
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96807
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-720 <STO>
A;Cross-references: UNIPROT:Q9CA22; GB:AE005173; NID:g6437532; PIDN:AAF08564.1; GSPDB:GN
C;Genetics:
A;Gene: T32E8.5
A;Map position: 1

  Query Match            11.0%;  Score 243.5;  DB 2;  Length 720;
  Best Local Similarity   25.9%;  Pred. No. 8.9e-07;
  Matches  106;  Conservative   56;  Mismatches  146;  Indels  101;  Gaps   17;

Qy         11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASAAE 70
                 :|  | :  |:|:|   ::||| |: |::|    :      |||:    |   | |
Db        331 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 383

Qy         71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
              ||| :|   |::|:|  ||: |          |  : || |:|: ::|
Db        384 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 436

Qy        121 ---LLLYS----SHQGCSMW------------MIQHCARDVLEALAFLHHEGYVHADLK 160
                 :| |     :|      |              ::   : :|:|:  :| |  ||:|||
Db        437 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 496

Qy        161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
              | | |       ||||||:: |  | |    || |       |:||| :|    :  |
Db        497 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 551

Qy        215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272
              :  :|    |:|||| || :|  | :          | |    | | |       : |
Db        552 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 607

Qy        273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH---------------------- 311
                | ||:|  |  | ::|       |  || : |  ||
Db        608 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 663

Qy        312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351
                      :| |:  | |: |      :| | : :: : ::  | | |
Db        664 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 706


RESULT 4
JC5153
mitogen-activated protein kinase (EC 2.7.1.-) - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: JC5153
R;Doerig, C.M.; Parzy, D.; Langsley, G.; Horrocks, P.; Carter, R.; Doerig, C.D.
Gene 177, 1-6, 1996
A;Title: A MAP kinase homologue from the human malaria parasite, Plasmodium falciparum.
A;Reference number: JC5153; MUID:97080491; PMID:8921836
A;Accession: JC5153
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-826 <DOE>
A;Cross-references: UNIPROT:Q94656; EMBL:X82646
C;Keywords: phosphotransferase

  Query Match            11.0%;  Score 243.5;  DB 2;  Length 826;
  Best Local Similarity   25.8%;  Pred. No. 1e-06;
  Matches  105;  Conservative   58;  Mismatches  145;  Indels   99;  Gaps   20;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASA--AEYGFRKERAAL 80
              : :  ::| |:   |:: ||  |       |:|:       ||  :   |:  || |   |
Db         23 YDILKKVGKGAYGVVFKGRCKKNKNIV--AVKKIF-----GAFQNCTDAQRTFR-EIIFL 74

Qy         81 EQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSEL----LLYSSHQGCSMWMIQ 136
              :| || ||: |  |     : |      |: : ::  : |:    ||   |:       :
Db         75 YELNGHDNIIKLMDVIK---AKNDNDIYLIFDFMETDLHEVIKADLLEEIHK-------K 124

Qy        137 HCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEGNQDV 189
              :   :| || ::|  | :| |:|| |||    :||  |: ||||:          |    :
Db        125 YIIYQLLRALKYIHSGGLLHRDIKPSNIL--VNSECHIKVADFGLARSISTHVNENKVPI 182

Qy        190 --KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL---KHT 244
                |: |  |||||           |   |  |  ||:|||| |: |:  |  |     |
Db        183 LTDYVATRWYRAPEI----------LLGSTHYTEDVDMWSLGCIMGELLCGKPLFTGNST 232

Qy        245 VRSQE------WKANSSAIID--HIFASKAVVN-AAIPAYHLR-----------DLIKSM 284
              :   |       | |   | |    || | : :   :   :|:           ||:: :
Db        233 MNQLEKIIQVIGKPNKKDIEDIRSPFAEKIISSFVDLKKKNLKDICYKASNESLDLLEKL 292

Qy        285 LHDDPSRRIPAEMAL----CSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEY 340
              |  :||:|| || ||       | ||    |    :: :|
Db        293 LQFNPSKRISAENALKHKYVEEFHSIIDEPTCRHIITIP---------------------- 331

Qy        341 EDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQK 387
                : :: |      |  ||  :: : |      |      | |:||  :|
```

```
Db        332 --INDNTKYRVNFYRNVVYFVIMRRNKFHSNVL----NQGESKKEEK 372


RESULT 5
T37758
protein kinase skp1p - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37758
R;Wedler, H.; Wambutt, R.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z21744
A;Accession: T37758
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-387 <WED>
A;Cross-references: UNIPROT:Q10452; EMBL:AL035064; PIDN:CAA22609.1; GSPDB:GN00066; SPDB:
A;Experimental source: strain 972h-; cosmid c1687
C;Genetics:
A;Gene: SPDB:SPAC1687.15
A;Map position: 1
A;Introns: 7/1; 70/3
C;Superfamily: kinase-related transforming protein; protein kinase homology

  Query Match            10.9%;  Score 241.5;  DB 2;  Length 387;
  Best Local Similarity   24.9%;  Pred. No. 6e-07;
  Matches   79;  Conservative   47;  Mismatches  112;  Indels   79;  Gaps   10;

Qy         29 LGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRN 88
              :||||   | :|    :       |:|: |          :  |:     : ::  | |
Db         38 VGSGSFGVVMQVHLIESDSK--AAIKRVL----------QDKRFKNRELQIMRIMKHPN 84

Qy         89 IVTLYGVF-------------TIHFSPNV---PSRCLLLELLDVSVSELLLYSSHQGCS 131
              || |   :             : | |      ||    : | : : |: ||
Db         85 IVDLIAYYYTTGDNSDEVYLNLVLEFMPETIYRASRLYTRQKLSMPMLEVKLY------- 137

Qy        132 MWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLS--FKEGNQDV 189
                      :| :||::|  |  | |:||:|:|   ||   || ||| :     |  :|
Db        138 -------IYQLLRSLAYIHASGICHRDIKPQNLLLDPENGILKLCDFGSAKILVAGEPNV 190

Qy        190 KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSG---------- 238
               || :  |||||          :  |: | |:|:|| | :: |:  |
Db        191 SYICSRYYRAPEL----------IFGATDYTHAIDIWSTGCVMAELMLGHPLFPGESGID 240

Qy        239 -----MKLKHTVRSQEWKANSSAIIDHIFAS------KAVVNAAIPAYHLRDLIKSMLHD 287
                  :|:  |   :: | :   ::| |           | : ::|   | ||:  ||
Db        241 QLVEIIKILGTPSREQIKTMNPNYMEHRFPQIRPQPLSRVFSRSVPLDAL-DLLSKMLQY 299

Qy        288 DPSRRIPAEMALCSPFF 304
               |: |: |  |:| |||
Db        300 TPTDRLTAAEAMCHPFF 316


RESULT 6
S48879
protein kinase SMK1 (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YP9499.10; protein YPR054w
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C;Accession: S48879; S54076
R;Krisak, L.; Strich, R.; Winters, R.S.; Hall, J.P.; Mallory, M.J.; Kreitzer, D.; Tuan,
Genes Dev. 8, 2151-2161, 1994
A;Title: SMK1, a developmentally regulated MAP kinase, is required for spore wall assemb
A;Reference number: S48879; MUID:95047362; PMID:7958885
A;Accession: S48879
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-388 <KRI>
A;Cross-references: UNIPROT:P41808; EMBL:L35047; NID:g538516; PIDN:AAB59325.1; PID:g5385
R;Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54059
A;Accession: S54076
A;Molecule type: DNA
A;Residues: 1-388 <BAD>
A;Cross-references: EMBL:Z49219; NID:g805025; PIDN:CAA89172.1; PID:g805035; MIPS:YPR054w
A;Experimental source: strain AB972
C;Genetics:
A;Gene: SGD:SMK1
A;Cross-references: SGD:S0006258; MIPS:YPR054w
A;Map position: 16R
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase; signal transduction
F;36-306/Domain: protein kinase homology <KIN>
F;44-52/Region: protein kinase ATP-binding motif

  Query Match            10.9%;  Score 240;  DB 2;  Length 388;
  Best Local Similarity   25.8%;  Pred. No. 7.3e-07;
  Matches  103;  Conservative   61;  Mismatches  133;  Indels  102;  Gaps   19;

Qy         21 RLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASAAEYG-------- 72
              |   | | ||:    :::       ||     : |||  |  |    |   |
Db         10 RAINVASNLGAPQQRTIFAKERISIPGY--YEIIQFLGKGAYGTVCSVKFKGRSPAARIA 67

Qy         73 -------FRKE---RAALEQL------QGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDV 116
                     | ||   : |: :|      :||:||| |  :  :  ||     |   ||:|
Db         68 VKKISNIFNKEILLKRAIRELKFMNFFKGHKNIVNLIDLEIVTSSPYDGLYC-YQELIDY 126

Qy        117 SVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLI 176
               ::: :::|| |  | : |::    :|  | ::|    :| |||| ||| :  | | |:
Db        127 DLAK-VIHSSVQ-LSEFHIKYFLYQILCGLKYIHSADVIHRDLKPGNILCTL-NGCLKIC 183

Qy        177 DFGLS------FKEGNQDVK-----YIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDL 225
              ||||:      | : :  |:     |: |  |||||           | |:   : :||:
Db        184 DFGLARGIHAGFFKCHSTVQPHITNYVATRWYRAPEL----------LLSNQPYSKSVDI 233

Qy        226 WSLGIILLE------MFSGMKLKHTV------------------RSQEW---KANSSAI 257
              |::| || |      :| |    | :                        : |   | ::: :
Db        234 WAVGCILAEFYARKPVFMGRDSMHQIFEIIKVLGTPDKDILIKFGTIKAWNLGKNSNNPV 293

Qy        258 IDHI-------FASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAP 310
                |       |||   :|           ||:|:|| | : |:  | |:  || :    |
Db        294 YKKIPWSNIFPFASHEAIN----------LIESLLHWDSTHRLNVEQAISHPFLNEVRKP 343

Qy        311 HIEDLVMLPTPVLRLLNVLDDDYLGNEEEYED-VVEDVK 348
                :: | |  |     :   :  | :  :  | :||:||
Db        344 D-DEPVCLQGP----FDFTYESELNSMSKLRDYLVEEVK 377


RESULT 7
S68680
stress-activated protein kinase (EC 2.7.1.-) 3 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S68680
R;Mertens, S.; Craxton, M.; Goedert, M.
FEBS Lett. 383, 273-276, 1996
A;Title: SAP kinase-3, a new member of the family of mammalian stress-activated protein k
A;Reference number: S68680; MUID:96198618; PMID:8925912
A;Accession: S68680
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-367 <MER>
A;Cross-references: UNIPROT:Q63538; EMBL:X96488; NID:g1262402; PIDN:CAA65342.1; PID:g1262
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;25-280/Domain: protein kinase homology <KIN>
F;33-41/Region: protein kinase ATP-binding motif

  Query Match            10.7%;  Score 235.5;  DB 2;  Length 367;
  Best Local Similarity   22.8%;  Pred. No. 1.2e-06;
  Matches   96;  Conservative   67;  Mismatches  127;  Indels  131;  Gaps   19;

Qy         10 AEPPRFLEAFGR------LWQVQS------RLGSGSSASVYRVRCCGNPGSPPG---ALK 54
                 : ||   : | |       |:|::       :|||:  :|     |    |  |   |:|
Db          2 SSPPPARKGFYRQEVTKTAWEVRAVYQDLQPVGSGAYGAV-----CSAVDSRTGNKVAIK 56
```

```
Qy          55 QFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRC------ 108
               :   |  :   |  |     :|   |: :: | |:: |  |||       |
Db          57 KLYRPFQSELFAKRA----YRELRLLKHMR-HENVIGLLDVFT-------PDETLDDFTD 104

Qy         109 --LLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILW 166
                 |::  :   : :|:    |:  |   ||     :|: | ::|  | :| |||| |:
Db         105 FYLVMPFMGTDLGKLM---KHETLSEDRIQFLVYQMLKGLKYIHAAGVIHRDLKPGNL-- 159

Qy         167 SAENEC-FKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDL 225
               :   :|  |::|||: :  ::   |: |  |||||  |          :    |  ||:
Db         160 AVNEDCELKILDFGLARQADSEMTGYVVTRWYRAPEVIL----------NWMRYTQTVDI 209

Qy         226 WSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNA-------------- 270
               ||:| |: || :|  |        :| |     ||:   | ::
Db         210 WSVGCIMAEMITGKIL--------FKGN-----DHLDQLKEIMKVTGTPPPEFVQKLQSA 256

Qy         271 -------AIPAYHLRD-------------LIKSMLHDDPSRRIPAEMALCSPFFSIPFA 309
                     :|    :|              |:: ||  |  :|: |  ||   |:|
Db         257 EAKNYMEGLPELEKKDFASVLTNASPQAVNLLEKMLVLDAEQRVTAAEALAHPYFE---- 312

Qy         310 PHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVK---EECQK--YGPVVSLLVPK 364
                           :: | :     ::|:|  :||    || ::  |  |:|   |:
Db         313 ----------------SLRDTEDEPKAQKYDDSFDDVDRTLEEWKRVTYKEVLSFKPPR 355

Qy         365 E 365
               :
Db         356 Q 356
```

RESULT 8
T45138
protein kinase skp1 [imported] - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T45138
R;Plyte, S.E.; Feoktistova, A.; Burke, J.D.; Woodgett, J.R.; Gould, K.L.
Mol. Cell. Biol. 16, 179-191, 1996
A;Title: Schizosaccharomyces pombe skp1+ encodes a protein kinase related to mammalian g
A;Reference number: Z22927; MUID:96104567; PMID:8524294
A;Accession: T45138
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-354 <PLY>
A;Cross-references: EMBL:L29449; NID:g1914882; PIDN:AAB51081.1; PID:g1914883
C;Genetics:
A;Gene: skp1
A;Introns: 6/3
C;Function:
A;Pathway: cells w
A;Note: cells with deletions in skp1 are sensitive to heat shock and exhibit defects in
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: phosphoprotein; phosphotransferase
F;335/Binding site: phosphate (Ser) (covalent) #status experimental

  Query Match   10.6%; Score 234.5; DB 2; Length 354;
  Best Local Similarity   24.6%; Pred. No. 1.4e-06;
  Matches   78; Conservative   47; Mismatches 113; Indels   79; Gaps   10;

```
Qy          29 LGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRN 88
               :||||   | :|    :      |:|: |          :  |:     : ::    |
Db          38 VGSGSFGVVMQVHLIESDSK--AAIKRVL----------QDKRFKNRELQIMRIMNDPN 84

Qy          89 IVTLYGVF-------------TIHFSPNV---PSRCLLLELLDVSVSELLLYSSHQGCS 131
               || |   :             : | |      ||    : | : |: ||
Db          85 IVDLIAYYYTTGDNSDEVYLNLVLEFMPETIYRASRLYTRQKLSMPMLEVKLY------- 137

Qy         132 MWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLS--FKEGNQDV 189
                       :| :||::|  |  | |:||:|:|   ||   || ||| :     |  :|
Db         138 -------IYQLLRSLAYIHASGICHRDIKPQNLLLDPENGILKLCDFGSAKILVAGEPNV 190

Qy         190 KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSG---------- 238
                || :  |||||           :   |: | |:|:|| | :: |:  |
```

```
Db         191 SYICSRYYRAPEL----------IFGATDYTHAIDIWSTGCVMAELMLGHPLFPGESGID 240

Qy         239 -----MKLKHTVRSQEWKANSSAIIDHIFAS------KAVVNAAIPAYHLRDLIKSMLHD 287
                    :|:  |   :: |  :   ::| |          | : ::|   | ||:  ||
Db         241 QLVEIIKILGTPSREQIKTMNPNYMEHRFPQIRPQPLSRVFSRSVPLDAL-DLLSKMLQY 299

Qy         288 DPSRRIPAEMALCSPFF 304
                |: |: |  |:| |||
Db         300 TPTDRLTAAEAMCHPFF 316
```

RESULT 9
S37790
probable serine/threonine-specific protein kinase (EC 2.7.1.-) YKL161c - yeast (Saccharom
N;Alternate names: protein YKL615
C;Species: Saccharomyces cerevisiae
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 16-Aug-2004
C;Accession: S37790; S37991; S44567
R;Vandenbol, M.; Bolle, P.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the EMBL Data Library, September 1993
A;Description: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4 loc
A;Reference number: S37786
A;Accession: S37790
A;Molecule type: DNA
A;Residues: 1-433 <VAN>
A;Cross-references: UNIPROT:P36005; EMBL:Z26877; NID:g407482; PIDN:CAA81493.1; PID:g4074
A;Experimental source: strain S288C
R;Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37976
A;Accession: S37991
A;Molecule type: DNA
A;Residues: 1-433 <VA2>
A;Cross-references: EMBL:Z28161; NID:g486280; PIDN:CAA82003.1; PID:g486281; MIPS:YKL161c
A;Experimental source: strain S288C
R;Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
Yeast 10, 35-40, 1994
A;Title: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4 loci of c
A;Reference number: S44563
A;Accession: S44567
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-433 <VA3>
A;Cross-references: EMBL:Z26877; NID:g407482; PIDN:CAA81493.1; PID:g407487
A;Experimental source: strain S288C
C;Genetics:
A;Cross-references: SGD:S0001644
A;Map position: 11L
A;Note: YKL161c
C;Superfamily: protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;21-287/Domain: protein kinase homology <KIN>
F;153/Active site: Asp #status predicted

  Query Match   10.6%; Score 234.5; DB 2; Length 433;
  Best Local Similarity   29.5%; Pred. No. 1.7e-06;
  Matches   89; Conservative   45; Mismatches   97; Indels   71; Gaps   16;

```
Qy          75 KERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLL--ELLDVSVSELLLYSSHQGCSM 132
               :|   | |:|| ||| |:    | | ||     : |  ||::  :|:::   | |
Db          71 RELKLLRHLRGHPNIVWLFDT-DIVFYPNGALNGVYLYEELMECDLSQII--RSEQRLED 127

Qy         133 WMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEG 185
                    |     :| || ::|    :| ||||:|:|    ::|  |: :||||       |
Db         128 AHFQSFIYQILCALKYIHSANVLHCDLKPKNLL--VNSDCQLKICNFGLSCSYSENHKVN 185

Qy         186 NQDVK-YIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLE------MFSG 238
               :  :| || :  |:|||           | :  ||| |||:|| | || |      || |
Db         186 DGFIKGYITSIWYKAPEI----------LLNYQECTKAVDIWSTGCILAELLGRKPMFEG 235

Qy         239 M----KLKHTVR---------SQEWKANSSAIIDHIF--------ASKAVVNAAIPAYHL 277
                  | | ::          ||     |  : ::||         : ::::  | |
Db         236 KDYVDHLNHILQILGTPPEETLQE--IASQKVYNYIFQFGNIPGRSFESILPGANP--EA 291
```

```
Qy        278 RDLIKSMLHDDPSRRIPAEMALCSPFFSI------PFA---------PHIEDLVMLPTPV 322
              :|:| ||  || :||  | ||  |: |:       |:          ||| :  |   |
Db        292 LELLKKMLEFDPKKRITVEDALEHPYLSMWHDIDEEFSCQKTFRFEFEHIESMAELGNEV 351

Qy        323 LR 324
              ::
Db        352 IK 353


RESULT 10
S64950
protein kinase HOG1 (EC 2.7.1.-), mitogen-activated - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein L2931; protein YLR113w
C;Species: Saccharomyces cerevisiae
C;Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 16-Aug-2004
C;Accession: S64950; S42145; S69400
R;Verhasselt, P.; Voet, M.; Volckaert, G.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64943
A;Accession: S64950
A;Molecule type: DNA
A;Residues: 1-435 <VER>
A;Cross-references: UNIPROT:P32485; EMBL:Z73285; NID:g1360507; PIDN:CAA97680.1; PID:e245
A;Experimental source: strain S288C
R;Brewster, J.L.; de Valoir, T.; Dwyer, N.D.; Winter, E.; Gustin, M.C.
Science 259, 1760-1763, 1993
A;Title: An osmosensing signal transduction pathway in yeast.
A;Reference number: S42145; MUID:93206121; PMID:7681220
A;Accession: S42145
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-408,'GQRSCSCK' <BRE>
A;Cross-references: EMBL:L06279; NID:g295610; PIDN:AAA34680.1; PID:g295611
R;Verhasselt, P.; Volckaert, G.
submitted to the EMBL Data Library, September 1995
A;Reference number: S69393
A;Accession: S69400
A;Molecule type: DNA
A;Residues: 1-435 <VEW>
A;Cross-references: EMBL:X89514; NID:g1297019; PIDN:CAA61691.1; PID:e198746; PID:g129702
C;Genetics:
A;Gene: SGD:HOG1; SSK3
A;Cross-references: SGD:S0004103; MIPS:YLR113w
A;Map position: 12R
C;Superfamily: protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase; signal tr
F;21-271/Domain: protein kinase homology <KIN>
F;29-37/Region: protein kinase ATP-binding motif
F;144/Active site: Asp #status predicted

  Query Match            10.6%;  Score 234;  DB 2;  Length 435;
  Best Local Similarity   25.7%;  Pred. No. 1.8e-06;
  Matches   98;  Conservative   53;  Mismatches  140;  Indels   90;  Gaps   18;

Qy         17 EAFGRLWQVQSR------LGSGSSASVYRVRCCGNP--GSPPGALKQFLPPGTTGAAASA 68
              : || :::: :|      :| |:    |    |     | | |:|: : | :|    |
Db         11 QIFGTVFEITNRYNDLNPVGMGAFGLV----CSATDTLTSQPVAIKKIMKPFSTAVLAKR 66

Qy         69 AEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQ 128
                    :|   |: |: | |:: |  :|    || :     : ||     :  ||     :
Db         67 T----YRELKLLKHLR-HENLICLQDIF---LSP-LEDIYFVTELQGTDLHRLL---QTR 114

Qy        129 GCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQD 188
                    :|:    :|  | ::|  | :| |||| ||| : ||   |: ||||:  :  |
Db        115 PLEKQFVQYFLYQILRGLKYVHSAGVIHRDLKPSNILIN-ENCDLKICDFGLARIQDPQM 173

Qy        189 VKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL----KHT 244
                |: |  |||||  |          : | |    ||:|| | |  ||  |  |      |
Db        174 TGYVSTRYYRAPEIML------TWQKYDVE----VDIWSAGCIFAEMIEGKPLFPGKDH- 222

Qy        245 VRSQEWKANSSAIIDHIFAS--KAVVNAAIPAYHLR---------------------- 278
                    :  :||  :  |  | |:|       |:

Db        223 -------VHQFSIITDLLGSPPKDVINTICSENTLKFVTSLPHRDPIPFSERFKTVEPDA 275

Qy        279 -DLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVM-----------LPTPVLRL 325
               ||:: ||  || :|| |  ||  | :| |:    :: |              ||    |:
Db        276 VDLLEKMLVFDPKKRITAADALAHP-YSAPYHDPTDEPVADAKFDWHFNDADLPVDTWRV 334

Qy        326 L---NVLDDDYLGNEEEYEDV 343
              :    :||   :|  :   |:
Db        335 MMYSEILDFHKIGGSDGQIDI 355


RESULT 11
A44878
protein kinase (EC 2.7.1.37) cdk2 [validated] - goldfish
C;Species: Carassius auratus (goldfish)
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: A44878
R;Hirai, T.; Yamashita, M.; Yoshikuni, M.; Tokumoto, T.; Kajiura, H.; Sakai, N.; Nagahama
Dev. Biol. 152, 113-120, 1992
A;Title: Isolation and characterization of goldfish cdk2, a cognate variant of the cell c
A;Reference number: A44878; MUID:92331802; PMID:1339336
A;Accession: A44878
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-298 <HIR>
A;Cross-references: UNIPROT:P43450; GB:S40289; NID:g251619; PIDN:AAB22550.1; PID:g251620
A;Experimental source: oocyte
A;Note: sequence extracted from NCBI backbone (NCBIN:108782, NCBIP:108783)
C;Genetics:
A;Gene: cdk2
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine/
F;2-255/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

  Query Match            10.6%;  Score 233;  DB 1;  Length 298;
  Best Local Similarity   28.7%;  Pred. No. 1.4e-06;
  Matches   99;  Conservative   48;  Mismatches  116;  Indels   82;  Gaps   17;

Qy         16 LEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQF-LPPGTTGAAASAAEYGFR 74
              :|:|     |   ::| |:   ||: :          |||:  |   | |  ::|
Db          1 MESF----QKVEKIGEGTYGVVYKAK--NKVTGETVALKKIRLDTETEGVPSTAI----- 49

Qy         75 KERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWM 134
              :| : |::|  | ||| |: |  ||    :    |: | |   :  :  |:  | |: :
Db         50 REISLLKEL-NHPNIVKLHDV--IHTENKL---YLVFEFLHQDLKRFMDSSTVTGISLPL 103

Qy        135 IQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY--- 191
              ::     :| ||| |    :| ||||:|:| :|: |  || ||||:   |      |
Db        104 VKSYLFQLLQGLAFCHSHRVLHRDLKPQNLLINAQGE-IKLADFGLARAFGVPVRTYTHE 162

Qy        192 IQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWK 251
              : |  |||||  | |          ::|||:|||| |  || :       |   :
Db        163 VVTLWYRAPEI-LLGC---------KYYSTAVDIWSLGCIFAEMIT--------RKALFP 204

Qy        252 ANSSAIIDHIFASKAVVN----------AAIPAYHL--------------------RDL 280
              :|   || :|     :              ::| |                      |||
Db        205 GDSE--IDQLFRIFRTLGTPDESIWPGVTSMPDYKPSFPKWARQDLSKVVPPLDEDGRDL 262

Qy        281 IKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRL 325
              :  ||  ||::|| |: ||   ||         |: | | | |||
Db        263 LGQMLIYDPNKRISAKNALVHRFF--------RDVTM-PVPPLRL 298


RESULT 12
A41227
protein kinase (EC 2.7.1.37) cdk2 - human
N;Alternate names: Eg1 homolog; protein kinase p34
C;Species: Homo sapiens (man)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C;Accession: A41227; S17873; S16520
R;Ninomiya-Tsuji, J.; Nomoto, S.; Yasuda, H.; Reed, S.I.; Matsumoto, K.
```

```
Proc. Natl. Acad. Sci. U.S.A. 88, 9006-9010, 1991
A;Title: Cloning of a human cDNA encoding a CDC2-related kinase by complementation of a
A;Reference number: A41227; MUID:92020980; PMID:1717994
A;Accession: A41227
A;Molecule type: mRNA
A;Residues: 1-298 <NIN>
A;Cross-references: UNIPROT:P24941; GB:M68520; NID:g180177; PIDN:AAA35667.1; PID:g180178
R;Tsai, L.H.; Harlow, E.; Meyerson, M.
Nature 353, 174-177, 1991
A;Title: Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E1A-
A;Reference number: S17873; MUID:91367262; PMID:1653904
A;Accession: S17873
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-298 <TSA>
A;Cross-references: GB:X62071; NID:g312802; PIDN:CAA43985.1; PID:g312803
R;Elledge, S.J.; Spottswood, M.R.
EMBO J. 10, 2653-2659, 1991
A;Title: A new human p34 protein kinase, CDK2, identified by complementation of a cdc28
A;Reference number: S16520; MUID:91330891; PMID:1714386
A;Accession: S16520
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-176,'S',178-298 <ELL>
A;Cross-references: EMBL:X61622; NID:g29848; PIDN:CAA43807.1; PID:g29849
C;Genetics:
A;Gene: GDB:CDK2
A;Cross-references: GDB:128984; OMIM:116953
A;Map position: 12q13-12q13
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine
F;2-255/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;14,160/Binding site: phosphate (Thr) (covalent) #status predicted
F;15/Binding site: phosphate (Tyr) (covalent) #status predicted
F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

  Query Match             10.6%;  Score 233;  DB 2;  Length 298;
  Best Local Similarity   29.0%;  Pred. No. 1.4e-06;
  Matches   98;  Conservative   43;  Mismatches  119;  Indels   78;  Gaps   16;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQF-LPPGTTGAAASAAEYGFRKERAALE 81
              :|   ::| |:   ||: |           |||:  |   | |  ::|      :| : |:
Db          4 FQKVEKIGEGTYGVVYKAR--NKLTGEVVALKKIRLDTETEGVPSTAI-----REISLLK 56

Qy         82 QLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARD 141
              :|  | ||| |  |  ||    :    |: | |  : : :  |:  |  : :|:
Db         57 EL-NHPNIVKLLDV--IHTENKL---YLVFEFLHQDLKKFMDASALTGIPLPLIKSYLFQ 110

Qy        142 VLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY---IQTDGYR 198
              :|: ||| |    :| ||||:|:| : |    || ||||:   |      |   : |  ||
Db        111 LLQGLAFCHSHRVLHRDLKPQNLLINTEG-AIKLADFGLARAFGVPVRTYTHEVVTLWYR 169

Qy        199 APEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAII 258
              |||  |  |            ::|||:|||| |  ||           |   :  :|   |
Db        170 APEI-LLGC---------KYYSTAVDIWSLGCIFAEM--------VTRRALFPGDSE--I 209

Qy        259 DHIFASKAVVN----------AAIPAYHL--------------------RDLIKSMLHD 287
              | :|    :           ::| |                       | |:  |||
Db        210 DQLFRIFRTLGTPDEVVWPGVTSMPDYKPSFPKWARQDFSKVVPPLDEDGRSLLSQMLHY 269

Qy        288 DPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRL 325
              ||::|| |: ||  |||         :| |  | | |||
Db        270 DPNKRISAKAALAHPFF--------QD-VTKPVPHLRL 298


RESULT 13
A56492
protein kinase ERK2 (EC 2.7.1.-) - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56492
R;Segall, J.E.; Kuspa, A.; Shaulsky, G.; Ecke, M.; Maeda, M.; Gaskins, C.; Firtel, R.A.;
J. Cell Biol. 128, 405-413, 1995
A;Title: A MAP kinase necessary for receptor-mediated activation of adenylyl cyclase in I
A;Reference number: A56492; MUID:95146547; PMID:7844154
A;Accession: A56492
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-369 <SEG>
A;Cross-references: UNIPROT:Q7M445; GB:L33043
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;12-273/Domain: protein kinase homology <KIN>
F;20-28/Region: protein kinase ATP-binding motif

  Query Match             10.6%;  Score 233;  DB 2;  Length 369;
  Best Local Similarity   26.9%;  Pred. No. 1.7e-06;
  Matches   91;  Conservative   52;  Mismatches  109;  Indels   86;  Gaps   17;

Qy         66 ASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSEL---- 121
              |: |:  || |   |::| || ||: |  |     : |     |: | ::  :  :
Db         52 ATDAQRTFR-EIMFLQELHGHENIIKLLNVIK---ADNDRDIYLVFEHMETDLHAVIRAK 107

Qy        122 LLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECF-KLIDFGL 180
              :|   |:       |:    :|:|| ::|    :| |:|| |:|    :||  |: ||||
Db        108 ILEEIHK-------QYTIYQLLKALKYMHSANVLHRDIKPSNLL--LNSECLVKVADFGL 158

Qy        181 --------SFKEGNQDV-KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGII 231
                      |  | |  : :|: |  |||||           |   |: |  ||:||:| |
Db        159 ARSITSLESIAEANPVLTEYVATRWYRAPEI----------LLGSTKYTKGVDMWSIGCI 208

Qy        232 LLE------MFSG----------MKLKHTVRSQEWKANSSAIIDHIFASKAVVN------ 269
              | |      || |          :::    ::: :|  |     :  |    |
Db        209 LGELLGEKAMFPGNSTMNQLDLIIEVTGRPSAEDIEAIKSPFAGTMLESLPPSNPRSLSD 268

Qy        270 ----AAIPAYHLRDLIKSMLHDDPSRRIPAEMALCSPF---FSIPF-APHIEDLVMLPT- 320
                  |:: |    ||:| :   :| :|| || ||   ||    |  |   || : :: :
Db        269 MYPSASVDAL---DLLKKLSQFNPDKRITAEEALAHPFVTQFHNPAEEPHFDRIIKISID 325

Qy        321 -----PVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQK 353
                  |:    | |           | |:::  ||| :|
Db        326 DGQKFPIAEYRNRL----------YNDIIKKKKEERKK 353


RESULT 14
T13024
probable protein kinase (EC 2.7.1.-) F8L21.120 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jan-2000
C;Accession: T13024
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mewe
submitted to the Protein Sequence Database, July 1999
A;Reference number: Z17587
A;Accession: T13024
A;Molecule type: DNA
A;Residues: 1-373 <BEV>
A;Cross-references: EMBL:AL096882; GSPDB:GN00062; ATSP:F8L21.120
A;Experimental source: cultivar Columbia; BAC clone F8L21
C;Genetics:
A;Gene: ATSP:F8L21.120
A;Map position: 4
A;Introns: 58/1; 98/3; 144/3; 255/3; 317/1
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: phosphotransferase; protein kinase
F;40-295/Domain: protein kinase homology <KIN>

  Query Match             10.4%;  Score 230;  DB 2;  Length 373;
  Best Local Similarity   25.9%;  Pred. No. 2.6e-06;
  Matches   83;  Conservative   59;  Mismatches  111;  Indels   68;  Gaps   14;

Qy         75 KERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWM 134
              :|   | |: | |:| :  :       :     :: ||:|  : :::   |:|  :
Db         86 REIKLLRHLE-HENVVVIKDIIRPPKKEDFVDVYIVFELMDTDLHQII--RSNQSLNDDH 142

Qy        135 IQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL--SFKEGNQDVKYI 192
```

```
              |:    :|  | ::|    :| |||| |:| :: |   |: ||||  :  |     :|:
Db        143 CQYFLYQILRGLKYIHSANVLHRDLKPSNLLLNS-NCDLKITDFGLARTTSETEYMTEYV 201

Qy        193 QTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLE------MFSGMKLKHTVR 246
              |  |||||           | : :| |||:|:||:| |  |      :| |    | ::
Db        202 VTRWYRAPEL----------LLNSSEYTSAIDVWSVGCIFAEIMTREPLFPGKDYVHQLK 251

Qy        247 ------------SQEW--KANSSAIIDHI-------FASK--AVVNAAIPAYHLRDLIKS 283
                        | |:   ||:   :  :       |:::  :: : ||      ||::
Db        252 LITELIGSPDGASLEFLRSANARKYVKELPKFPRQNFSARFPSMNSTAI------DLLEK 305

Qy        284 MLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDV 343
              ||  || :||  | ||| |: |      : ||   |        |  : :  : |:
Db        306 MLVFDPVKRITVEEALCYPYLSA-----LHDLNDEP--------VCSNHFSFHFEDPSST 352

Qy        344 VEDVKE----ECQKYGPVVSL 360
               |::||    |  |: |: |:
Db        353 EEEIKELVWLESVKFNPLPSI 373


RESULT 15
S40471
mitogen-activated protein kinase 5 (EC 2.7.1.-) - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 07-Sep-1994 #sequence_revision 19-Jan-1996 #text_change 09-Jul-2004
C;Accession: S40471
R;Mizoguchi, T.; Hayashida, N.; Yamaguchi-Shinozaki, K.; Kamada, H.; Shinozaki, K.
FEBS Lett. 336, 440-444, 1993
A;Title: ATMPKs: a gene family of plant MAP kinases in Arabidopsis thaliana.
A;Reference number: S40469; MUID:94109583; PMID:8282107
A;Accession: S40471
A;Molecule type: mRNA
A;Residues: 1-376 <MIZ>
A;Cross-references: UNIPROT:Q39025; EMBL:D21841; NID:g457401; PIDN:BAA04868.1; PID:g4574
C;Genetics:
A;Gene: MPK5
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;40-298/Domain: protein kinase homology <KIN>
F;49-57/Region: protein kinase ATP-binding motif

  Query Match             10.4%;  Score 230;  DB 2;  Length 376;
  Best Local Similarity   25.9%;  Pred. No. 2.6e-06;
  Matches   83;  Conservative   58;  Mismatches  112;  Indels   68;  Gaps   14;

Qy         75 KERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWM 134
              :|   |  |: | |:| :  :       :     :: ||:|  : :::   |:|  :
Db         89 REIKLLRHLE-HENVVVIKDIIRPPKKEDFVDVYIVFELMDTDLHQII--RSNQSLNDDH 145

Qy        135 IQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL--SFKEGNQDVKYI 192
              |:    :|  | ::|    :| |||| |:| :: |   |: ||||  :  |     :|:
Db        146 CQYFLYQILRGLKYIHSANVLHRDLKPSNLLLNS-NCDLKITDFGLARTTSETEYMTEYV 204

Qy        193 QTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLE------MFSGMKLKHTVR 246
              |  |||||           | : :| |||:|:||:| |  |      :| |    | ::
Db        205 VTRWYRAPEL----------LLNSSEYTSAIDVWSVGCIFAEIMTREPLFPGKDYVHQLK 254

Qy        247 ------------SQEW--KANSSAIIDHI-------FASK--AVVNAAIPAYHLRDLIKS 283
                        | |:   ||    :  :       |:::  :: : ||      ||::
Db        255 LITELIGSPDGASLEFLRSANGGKYVKELPKFPRQNFSARFPSMNSTAI------DLLEK 308

Qy        284 MLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDV 343
              ||  || :||  | ||| |: |      : ||   |        |  : :  : |:
Db        309 MLVFDPVKRITVEEALCYPYLSA-----LHDLNDEP--------VCSNHFSFHFEDPSST 355

Qy        344 VEDVKE----ECQKYGPVVSL 360
               |::||    |  |: |: |:
Db        356 EEEIKELVWLESVKFNPLPSI 376


Search completed: November 29, 2004, 11:42:40
Job time : 28 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 122.5 Seconds
(without alignments)
1968.016 Million cell updates/sec

Title: US-10-798-532-2
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAFG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2198	99.6	419	1	KIST_HUMAN	Q8tas1 homo sapien
2	2186	99.1	419	1	KIST_RAT	Q63285 rattus norv
3	2185	99.0	419	2	AAH58732	Aah58732 mus muscu
4	2185	99.0	419	2	AAO13515	Aao13515 mus muscu
5	2180	98.8	419	1	KIST_MOUSE	P97343 mus musculu
6	1711	77.6	330	2	Q8C0N6	Q8c0n6 mus musculu
7	257	11.7	505	2	Q6CFX3	Q6cfx3 yarrowia li
8	248.5	11.3	882	2	Q8WQL7	Q8wql7 caenorhabdi
9	243.5	11.0	395	2	Q6GM90	Q6gm90 xenopus lae
10	243.5	11.0	720	2	Q9CA22	Q9ca22 arabidopsis
11	243.5	11.0	765	2	Q25758	Q25758 plasmodium
12	243.5	11.0	777	2	Q84VX4	Q84vx4 arabidopsis
13	243.5	11.0	826	2	Q94656	Q94656 plasmodium
14	243.5	11.0	914	2	Q8ILF0	Q8ilf0 plasmodium
15	242.5	11.0	395	2	Q8AYR3	Q8ayr3 xenopus lae
16	242.5	11.0	395	2	AAH70745	Aah70745 xenopus l
17	242	11.0	433	2	Q84SN3	Q84sn3 oryza sativ
18	241.5	10.9	387	1	SKP1_SCHPO	Q10452 schizosacch
19	240.5	10.9	360	2	Q6SZ67	Q6sz67 lytechinus
20	240.5	10.9	360	2	AAR17088	Aar17088 lytechinu
21	240	10.9	388	1	SMK1_YEAST	P41808 saccharomyc
22	238	10.8	298	1	CDK2_MESAU	P48963 mesocricetu
23	235.5	10.7	367	1	MK12_RAT	Q63538 rattus norv
24	235.5	10.7	434	2	Q75B85	Q75b85 ashbya goss
25	235.5	10.7	434	2	AAS51605	Aas51605 ashbya go
26	234.5	10.6	367	1	MK12_MOUSE	O08911 mus musculu
27	234.5	10.6	392	2	Q6ZD93	Q6zd93 oryza sativ
28	234.5	10.6	392	2	BAC99508	Bac99508 oryza sat
29	234.5	10.6	433	1	KKQ1_YEAST	P36005 saccharomyc
30	234	10.6	298	1	CDK2_CRIGR	O55076 cricetulus
31	234	10.6	298	1	CDK2_RAT	Q63699 rattus norv
32	234	10.6	298	2	Q6P751	Q6p751 rattus norv
33	234	10.6	298	2	AAH61832	Aah61832 rattus no
34	234	10.6	361	1	JNK_SUBDO	Q966y3 suberites d
35	234	10.6	435	1	HOG1_YEAST	P32485 saccharomyc
36	233.5	10.6	433	2	Q6BG29	Q6bg29 paramecium
37	233.5	10.6	1233	2	Q6CAC2	Q6cac2 yarrowia li
38	233	10.6	298	1	CDK2_CARAU	P43450 carassius a
39	233	10.6	298	1	CDK2_HUMAN	P24941 homo sapien
40	233	10.6	298	2	AAP35467	Aap35467 homo sapi
41	233	10.6	369	2	Q7M445	Q7m445 dictyosteli
42	233	10.6	392	2	Q9VMN3	Q9vmn3 drosophila
43	232.5	10.5	357	2	Q9HD31	Q9hd31 homo sapien
44	232.5	10.5	385	2	Q8IU85	Q8iu85 homo sapien
45	232.5	10.5	385	2	Q8BW96	Q8bw96 mus musculu

ALIGNMENTS

RESULT 1
KIST_HUMAN
ID KIST_HUMAN STANDARD; PRT; 419 AA.
AC Q8TAS1; Q96C22;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting
DE with stathmin).
GN Name=KIST; Synonyms=KIS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX PubMed=12782393;
RA Bieche I., Manceau V., Curmi P.A., Laurendeau I., Lachkar S.,
RA Leroy K., Vidaud D., Sobel A., Maucuer A.;
RT "Quantitative RT-PCR reveals a ubiquitous but preferentially neural
RT expression of the KIS gene in rat and human.";
RL Brain Res. Mol. Brain Res. 114:55-64(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Eye, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Possesses kinase activity. May be involved in
CC trafficking and/or processing of RNA (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Interacts with stathmin (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;

```
CC         IsoId=Q8TAS1-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q8TAS1-2; Sequence=VSP_004908, VSP_004909;
CC         Note=No experimental confirmation available;
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC   -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ536197; CAD60192.1; -.
DR   EMBL; BC014917; AAH14917.1; -.
DR   EMBL; BC026046; AAH26046.1; -.
DR   HSSP; P26368; 1OOP.
DR   GO; GO:0005737; C:cytoplasm; ISS.
DR   GO; GO:0005634; C:nucleus; ISS.
DR   GO; GO:0005524; F:ATP binding; ISS.
DR   GO; GO:0005515; F:protein binding; ISS.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
DR   GO; GO:0003723; F:RNA binding; ISS.
DR   GO; GO:0018105; P:peptidyl-serine phosphorylation; ISS.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   Pfam; PF00076; RRM_1; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00360; RRM; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR   PROSITE; PS50102; RRM; 1.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW   Alternative splicing; ATP-binding; Nuclear protein; RNA-binding;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       23    304       Protein kinase.
FT   DOMAIN      324    406       RNA-binding (RRM).
FT   NP_BIND      29     37       ATP (By similarity).
FT   BINDING      54     54       ATP (By similarity).
FT   ACT_SITE    158    158       Proton acceptor (By similarity).
FT   VARSPLIC    342    344       DVV -> GLC (in isoform 2).
FT                                /FTId=VSP_004908.
FT   VARSPLIC    345    419       Missing (In isoform 2).
FT                                /FTId=VSP_004909.
SQ   SEQUENCE   419 AA;  46546 MW;  903E982EE12A8CF8 CRC64;

  Query Match             99.6%;  Score 2198;  DB 1;  Length 419;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-159;
  Matches  418;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEEEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 2
KIST_RAT
ID   KIST_RAT       STANDARD;      PRT;   419 AA.
AC   Q63285;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting
DE   with stathmin) (PAM COOH-terminal interactor protein 2) (P-CIP2).
GN   Name=Kist; Synonyms=Kis;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, TISSUE
RP   SPECIFICITY, AND MUTAGENESIS OF LYS-54.
RC   TISSUE=Striatum;
RX   MEDLINE=97435279; PubMed=9287318;
RA   Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P.,
RA   Sobel A.;
RT   "KIS is a protein kinase with an RNA recognition motif.";
RL   J. Biol. Chem. 272:23151-23156(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Caldwell B.D., Darlington D.N., Penzes P., Johnson R.C., Eipper B.A.,
RA   Mains R.E.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 28-419 FROM N.A.
RC   TISSUE=Hippocampus;
RX   MEDLINE=97067094; PubMed=8910496;
RA   Alam R., Caldwell B.D., Johnson R.C., Darlington D.N., Mains R.E.,
RA   Eipper B.A.;
RT   "Novel proteins that interact with the COOH-terminal cytosolic routing
RT   determinants of an integral membrane peptide-processing enzyme.";
RL   J. Biol. Chem. 271:28636-28640(1996).
CC   -!- FUNCTION: Possesses kinase activity. May be involved in
CC       trafficking and/or processing of RNA.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBUNIT: Interacts with stathmin (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC   -!- TISSUE SPECIFICITY: In the embryo, preferentially expressed in the
CC       developing nervous system.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC   -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X98374; CAA67021.1; -.
DR   EMBL; U70372; AAC53031.2; -.
DR   HSSP; P26368; 1OOP.
DR   RGD; 2968; Kist.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
```

```
DR   GO; GO:0005524; F:ATP binding; IC.
DR   GO; GO:0005515; F:protein binding; ISS.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
DR   GO; GO:0003723; F:RNA binding; NAS.
DR   GO; GO:0018105; P:peptidyl-serine phosphorylation; IDA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   Pfam; PF00076; RRM_1; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00360; RRM; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR   PROSITE; PS50102; RRM; 1.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW   ATP-binding; Nuclear protein; RNA-binding;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       23    304       Protein kinase.
FT   DOMAIN      324    406       RNA-binding (RRM).
FT   NP_BIND      29     37       ATP (By similarity).
FT   BINDING      54     54       ATP.
FT   ACT_SITE    141    141       Proton acceptor (By similarity).
FT   ACT_SITE    158    158       Proton acceptor (By similarity).
FT   MUTAGEN      54     54       K->R: Loss of activity.
SQ   SEQUENCE   419 AA;  46505 MW;  D2F9B7BF8F080DF3 CRC64;

  Query Match             99.1%;  Score 2186;  DB 1;  Length 419;
  Best Local Similarity   99.0%;  Pred. No. 1.1e-158;
  Matches  415;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGTPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||||||| ||:|||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 3
AAH58732
ID   AAH58732    PRELIMINARY;      PRT;   419 AA.
AC   AAH58732;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Kinase interacting with leukemia-associated gene (Stathmin).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC058732; AAH58732.1; -.
KW   Kinase.
SQ   SEQUENCE   419 AA;  46489 MW;  3BD3C06A59E22540 CRC64;

  Query Match             99.0%;  Score 2185;  DB 2;  Length 419;
  Best Local Similarity   99.0%;  Pred. No. 1.3e-158;
  Matches  415;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGTPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||||||| ||:|||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 4
AAO13515
ID   AAO13515    PRELIMINARY;      PRT;   419 AA.
AC   AAO13515;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
```

```
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   KIS kinase.
GN   KIS.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129S6;
RA   Crook M.F., Boehm M., Nabel E.G.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY180177; AAO13515.1; -.
KW   Kinase.
SQ   SEQUENCE   419 AA;  46489 MW;  3BD3C06A59E22540 CRC64;

  Query Match             99.0%;  Score 2185;  DB 2;  Length 419;
  Best Local Similarity   99.0%;  Pred. No. 1.3e-158;
  Matches  415;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGTPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||||| ||:|||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 5
KIST_MOUSE
ID   KIST_MOUSE     STANDARD;      PRT;   419 AA.
AC   P97343; Q61775; Q9CYT1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting
DE   with stathmin).
GN   Name=Kist; Synonyms=KIS;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [2]
RP   SEQUENCE OF 1-414 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=97435279; PubMed=9287318;
RA   Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P.,
RA   Sobel A.;
RT   "KIS is a protein kinase with an RNA recognition motif.";
RL   J. Biol. Chem. 272:23151-23156(1997).
RN   [3]
RP   REVISIONS TO 52 AND 68-69.
RA   Maucuer A.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 122-321 FROM N.A., AND INTERACTION WITH STATHMIN.
RC   TISSUE=Embryo;
RX   MEDLINE=95241452; PubMed=7724523;
RA   Maucuer A., Camonis J.H., Sobel A.;
RT   "Stathmin interaction with a putative kinase and coiled-coil-forming
RT   protein domains.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:3100-3104(1995).
CC   -!- FUNCTION: Possesses kinase activity. May be involved in
CC       trafficking and/or processing of RNA (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBUNIT: Interacts with stathmin.
CC   -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC   -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Y10725; CAA71714.2; -.
DR   EMBL; AK013347; BAB28802.1; -.
DR   EMBL; X82320; CAA57763.1; -.
DR   PIR; I48615; I48615.
DR   HSSP; P26368; 1OOP.
DR   MGD; MGI:1341908; Kist.
DR   GO; GO:0005737; C:cytoplasm; ISS.
DR   GO; GO:0005634; C:nucleus; ISS.
DR   GO; GO:0005524; F:ATP binding; ISS.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
DR   GO; GO:0003723; F:RNA binding; ISS.
DR   GO; GO:0018105; P:peptidyl-serine phosphorylation; ISS.
DR   InterPro; IPR011009; Kinase_like.
```

```
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   Pfam; PF00076; RRM_1; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00360; RRM; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR   PROSITE; PS50102; RRM; 1.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW   ATP-binding; Nuclear protein; RNA-binding;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       23    303       Protein kinase.
FT   DOMAIN      323    405       RNA-binding (RRM).
FT   NP_BIND      29     37       ATP (By similarity).
FT   BINDING      54     54       ATP (By similarity).
FT   ACT_SITE    158    158       Proton acceptor (By similarity).
FT   CONFLICT    170    170       N -> D (in Ref. 4).
FT   CONFLICT    231    231       T -> I (in Ref. 3 and 4).
SQ   SEQUENCE   419 AA;  46477 MW;  A91B307BA50F29A1 CRC64;

  Query Match             98.8%;  Score 2180;  DB 1;  Length 419;
  Best Local Similarity   98.8%;  Pred. No. 3.2e-158;
  Matches  414;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGTPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGITLLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||||| ||:|||||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 6
Q8C0N6
ID   Q8C0N6      PRELIMINARY;      PRT;   330 AA.
AC   Q8C0N6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE   library, clone:4933401D07 product:kinase interacting with leukemia-
DE   associated gene (stathmin), full insert sequence.
GN   Name=Kist;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=99279253; PubMed=10349636;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=20499374; PubMed=11042159;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=20530913; PubMed=11076861;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK030152; BAC26812.1; -.
DR   HSSP; P26368; 1O0P.
DR   MGD; MGI:1341908; Kist.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004672; F:protein kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00069; Pkinase; 2.
DR   Pfam; PF00076; RRM_1; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00360; RRM; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
```

```
DR   PROSITE; PS50102; RRM; 1.
KW   Kinase.
SQ   SEQUENCE   330 AA;  37201 MW;  4164C0758EE7CB4C CRC64;

  Query Match             77.6%;  Score 1711;  DB 2;  Length 330;
  Best Local Similarity   98.8%;  Pred. No. 1.7e-122;
  Matches  326;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         90 VTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFL 149
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFL 60

Qy        150 HHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLA 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLA 120

Qy        210 QAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVN 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVN 180

Qy        270 AAIPAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVL 329
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 AAIPAYHLRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVL 240

Qy        330 DDDYLGNEEEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLL 389
              ||||| ||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DDDYLENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLL 300

Qy        390 TGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||
Db        301 TGRMFDGKFVVATFYPLSAYKRGYLYQTLL 330


RESULT 7
Q6CFX3
ID   Q6CFX3      PRELIMINARY;      PRT;   505 AA.
AC   Q6CFX3;
DT   01-OCT-2004 (TrEMBLrel. 28, Created)
DT   01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Chromosome B of strain CLIB99 of Yarrowia lipolytica.
GN   ORFNames=YALI0B02816g;
OS   Yarrowia lipolytica (Candida lipolytica).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Dipodascaceae; Yarrowia.
OX   NCBI_TaxID=4952;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIB99;
RG   GENOLEVURES;
RA   Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA   Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA   Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA   Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA   Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA   Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA   Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA   Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA   Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA   Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA   Swennene D., Tekaia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA   Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA   Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA   Wincker P., Souciet J.L.;
RT   "Genome evolution in yeasts.";
RL   Nature 430:35-44(2004).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIB99;
RA   Genoscope;
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR382128; CAG82657.1; -.
SQ   SEQUENCE   505 AA;  58191 MW;  D30933013FCEE357 CRC64;

  Query Match             11.7%;  Score 257;  DB 2;  Length 505;
  Best Local Similarity   25.5%;  Pred. No. 4.5e-11;
  Matches   88;  Conservative   46;  Mismatches  129;  Indels   82;  Gaps   11;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQ 82
              :||   || |:   |   :  | |  |  |:|:     :        |     :|   |
Db         22 FQVTKELGHGAYGIVCAAKYTGTPDGPGVAIKKVTNIFSKNILCKRA----LRELKLLNH 77

Qy         83 LQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDV-SVSELLLYSSHQGCSMWMI------ 135
               :||:||  ||                ::::|  : :|| ||      | |  |
Db         78 FRGHKNITCLYD---------------MDIVDTNNFNELYLYEELMECDMHQIIRSGQP 121

Qy        136 ------QHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDV 189
                   |     :|  : ::|    :| |||| |:| :|: |  |: ||||:    : |
Db        122 LTDAHYQSFVYQILAGVKYIHSADVLHRDLKPGNLLVNADCE-LKICDFGLARGFSSDDE 180

Qy        190 K-------YIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLK 242
              |       |: |  |||||  |           |     | |:|:||:| || |:  |   |
Db        181 KNAGFLTEYVATRWYRAPEIML----------SFQSYTKAIDIWSVGCILAELLGG---K 227

Qy        243 HTVRSQEWKANSSAIIDHI----------------------FASKAVVNAAIPAYH-- 276
                 : : :    : |: ::                     |  |  :  |  :
Db        228 PLFKGKNYVDQLNQILHYLGTPSEETLKRIGSPRAQEYVRGLPFMPKIPFSTLFPTANPE 287

Qy        277 LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTP 321
               ||:: ||  ||: |:  | ||  |:  |    |   |  : |||
Db        288 ALDLLEKMLAFDPAERVTVEEALEHPYLKIWHDP--RDEPVCPTP 330


RESULT 8
Q8WQL7
ID   Q8WQL7      PRELIMINARY;      PRT;   882 AA.
AC   Q8WQL7; Q22155;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Minibrain kinase (Hypothetical protein T04C10.1).
GN   Name=mbk-1; Synonyms=T04C10.1;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22505392; PubMed=12618396;
RA   Raich W.B., Moorman C., Lacefield C.O., Lehrer J., Bartsch D.,
RA   Plasterk R.H., Kandel E.R., Hobert O.;
RT   "Characterization of Caenorhabditis elegans homologs of the Down
RT   syndrome candidate gene DYRK1A.";
RL   Genetics 163:571-580(2003).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Burton J.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY064464; AAL40874.1; -.
DR   EMBL; Z69885; CAA93756.2; -.
DR   HSSP; P24941; 1H0V.
DR   IntAct; Q8WQL7; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
```

```
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Hypothetical protein; Kinase;
KW   Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   882 AA;  101244 MW;  DD5C9DD354DA7098 CRC64;

  Query Match             11.3%;  Score 248.5;  DB 2;  Length 882;
  Best Local Similarity   25.7%;  Pred. No. 4e-10;
  Matches  104;  Conservative   57;  Mismatches  113;  Indels  131;  Gaps   17;

Qy         88 NIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALA 147
              ||||| | | :| :      ||: |||  :: :||  :| :| |: : :  |: : : |
Db        389 NIVTLKGHF-VHRA----HLCLVFELLSYNLYDLLKNTSFRGVSLNLARKFAQQLGKTLL 443

Qy        148 FLHHE--GYVHADLKPRNILW-SAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAEL 204
              ||       :| |||| |:|  :|:    ::|||| | : |::  :|||:  ||:||  |
Db        444 FLSSPELSIIHCDLKPENVLLVNAKRSQIRVIDFGSSCQTGHRIYQYIQSRFYRSPEVLL 503

Qy        205 QNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSG-------------MKL---------- 241
                    |:  ||:    :|:|||| ||:|| :|              ||:
Db        504 -------GIAYDTK----IDMWSLGCILVEMHTGEPLFAGSSEVDQMMKIVEVLGMPPKE 552

Qy        242 ----------------------------KHTVRSQEWKANSSAIIDHIFASKAVVNAAI 272
                                         :||     :||  :  :  |      | :
Db        553 MLDIGPKTHKYFDKTEDGIYYCKKTRDGYRHT-----YKAPGARKLHEILG----VTSGG 603

Qy        273 PA--------------YHLRDLIKSMLHDDPSRRIPAEMALCSPFFSI--PFAPHIEDLV 316
              |                    :|||| ||  || :||       ||:: :  ||    |: |
Db        604 PGGRRLGEPGHSVEDYSKFKDLIKRMLQFDPKQRI-------SPYYVVRHPFLKQKEERV 656

Qy        317 MLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKY--GPVVSLLVPKENPGRGQVFV 374
                 ||                        ::::: | |   |   :    |:|  | |:|
Db        657 PSQPPVSH-------------------SNLQQQQQLYIQQPSSQMSQVMESPSVGSVYV 696

Qy        375 E----YANAGDSKAAQKLLTGRMFDGKF--VVATFYPLSAYKRGY 413
              |    |  | | |    :|    :|    | |      ||::: |
Db        697 EDNGMYRQAPGSSANPISVTSSFDEGDAMEVDAGRRRFSAHQQNY 741


RESULT 9
Q6GM90
ID   Q6GM90      PRELIMINARY;      PRT;   395 AA.
AC   Q6GM90;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22341132; PubMed=12454917;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Klein S., Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC074183; AAH74183.1; -.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Hypothetical protein; Kinase;
KW   Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   395 AA;  43896 MW;  4C78B8BF828A2CC4 CRC64;

  Query Match             11.0%;  Score 243.5;  DB 2;  Length 395;
  Best Local Similarity   27.7%;  Pred. No. 3.5e-10;
  Matches   91;  Conservative   47;  Mismatches  108;  Indels   83;  Gaps   16;

Qy          4 SGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASV----------YRVRCCGNPGSPPGA 52
              ||  |  :    :|   |::| :  ||:|: : |          : |:|      |  |
Db          9 SGACWKKQ----VEDIKRMFQFKEVLGTGAFSEVVLAEEKETGKLFAVKCI-----PKKA 59

Qy         53 LKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLL- 111
              ||                |      | | ::: | ||| |  ::         ||   |:
Db         60 LK------------GKESSIENEIAVLRKIK-HENIVALEDIY------ESPSHLYLVM 99

Qy        112 ------ELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNIL 165
                   || |  : |   |:     ::       | ||:|:::||  | || |||| |:|
Db        100 QLVSGGELFD-RIVEKGFYTEKDASTL------IRQVLDAVSYLHRLGIVHRDLKPENLL 152

Qy        166 WSAENECFKLI--DFGLSFKEGNQDVKYIQ--TDGYRAPEAELQNCLAQAGLQSDTECTS 221
              : :: |  |::  |||||  ||  ||       | || |||       |||         :
Db        153 YFSQVEESKIMISDFGLSKMEGKGDVMSTACGTPGYVAPEV-----LAQ------KPYSK 201

Qy        222 AVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAY-----H 276
              ||| ||:|:|   :  |           :  | | : : |   ||       | :
Db        202 AVDCWSIGVIAYILLCGY-------PPFYDENDSRLFEQIL--KADYEFDSPYWDDISES 252

Qy        277 LRDLIKSMLHDDPSRRIPAEMALCSPFFS 305
               :| |:::: ||::|   | ||  |: :
Db        253 AKDFIQNLMEKDPNKRYTCEQALRHPWIA 281


RESULT 10
```

```
Q9CA22
ID   Q9CA22      PRELIMINARY;      PRT;   720 AA.
AC   Q9CA22;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein T32E8.5.
GN   Name=T32E8.5;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA   Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA   Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Town C.D., Kaul S.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC012193; AAG51619.1; -.
DR   PIR; A96807; A96807.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW   ATP-binding; Hypothetical protein; Kinase; Transferase.
SQ   SEQUENCE   720 AA;  80027 MW;  F26A9A0B1C6D91DE CRC64;

  Query Match             11.0%;  Score 243.5;  DB 2;  Length 720;
  Best Local Similarity   25.9%;  Pred. No. 7.5e-10;
  Matches  106;  Conservative   56;  Mismatches  146;  Indels  101;  Gaps   17;

Qy         11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASAAE 70
              :|  | :  |:|:|   ::||| |: |::|    :      |||:    |  | |
Db        331 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 383

Qy         71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
              ||| :|   |::|:|  ||: |            |  : || |:|: ::|
Db        384 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 436

Qy        121 ---LLLYS----SHQGCSMW------------MIQHCARDVLEALAFLHHEGYVHADLK 160
                :| |     :|     |              ::   : :|:|:  :| |  ||:|||
Db        437 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 496

Qy        161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
              | | |       ||||||:: |  | |   || |       | :||| : |   : |
Db        497 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 551

Qy        215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272
              :  :|    |:|||| || :|  | :          | |     | |  |       : |
Db        552 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 607

Qy        273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH---------------------- 311
                | ||:|  |  | ::|      |  || : |  ||
Db        608 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 663

Qy        312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351
                      :| |:  | |: |      :| | : :: : ::  | | |
Db        664 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 706
```

```
RESULT 11
Q25758
ID   Q25758      PRELIMINARY;      PRT;   765 AA.
AC   Q25758;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mitogen-activated protein kinase-related protein.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Dd2;
RA   Lin D.T., Goldman N.D., Syin C.;
RT   "Stage-specific expression of a Plasmodium falciparum protein related
RT   to the eukaryotic mitogen-activated protein kinases.";
RL   Mol. Biochem. Parasitol. 78:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Dd2;
RA   Lin D.T., Goldman N.D., Syin C.;
RL   Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; U36377; AAC47170.1; -.
DR   HSSP; Q16539; 1KV1.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR003527; MAP_kin.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS01351; MAPK; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   765 AA;  90231 MW;  5BEC8EA6A4B02D32 CRC64;

  Query Match             11.0%;  Score 243.5;  DB 2;  Length 765;
  Best Local Similarity   25.8%;  Pred. No. 8.1e-10;
  Matches  105;  Conservative   58;  Mismatches  145;  Indels   99;  Gaps   20;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASA--AEYGFRKERAAL 80
              :  ::| |:   |:: ||  |      |:|:       ||  :   |:  || |   |
Db         23 YDILKKVGKGAYGVVFKGRCKKNKNIV--AVKKIF-----GAFQNCTDAQRTFR-EIIFL 74

Qy         81 EQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSEL----LLYSSHQGCSMWMIQ 136
              :| || ||: |  |     : |      |: : ::  : |:    ||   |:       :
Db         75 YELNGHDNIIKLMDVIK---AKNDNDIYLIFDFMETDLHEVIKADLLEEIHK-------K 124

Qy        137 HCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEGNQDV 189
              :    :| || ::|  | :| |:|| |||    :||  |: ||||:        |    :
Db        125 YIIYQLLRALKYIHSGGLLHRDIKPSNIL--VNSECHIKVADFGLARSISTHVNENKVPI 182

Qy        190 --KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL---KHT 244
                |: |  |||||          |  |  |  ||:|||| |: |:  |  |      |
Db        183 LTDYVATRWYRAPEI----------LLGSTHYTEDVDMWSLGCIMGELLCGKPLFTGNST 232

Qy        245 VRSQE------WKANSSAIID--HIFASKAVVN-AAIPAYHLR----------DLIKSM 284
              :  |       | |   | |    || | : :   :   :|:           ||:: :
Db        233 MNQLEKIIQVIGKPNKKDIEDIRSPFAEKIISSFVDLKKKNLKDICYKASNESLDLLEKL 292

Qy        285 LHDDPSRRIPAEMAL----CSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEY 340
              |  :||:|| || ||       | ||   |    :: :|
```

```
Db        293 LQFNPSKRISAENALKHKYVEEFHSIIDEPTCRHIITIP---------- 331

Qy        341 EDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQK 387
Db        332 --INDNTKYRVNFYRNVVYFVIMRRNKFHSNVL----NQGESKKEEK 372


RESULT 12
Q84VX4
ID   Q84VX4      PRELIMINARY;      PRT;   777 AA.
AC   Q84VX4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   At1g77720.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA   Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA   Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA   Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA   Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA   Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA   Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT004636; AAO42882.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW   ATP-binding; Kinase; Transferase.
SQ   SEQUENCE   777 AA;  86376 MW;  F2C25DA9609A878F CRC64;

  Query Match           11.0%;  Score 243.5;  DB 2;  Length 777;
  Best Local Similarity   25.9%;  Pred. No. 8.2e-10;
  Matches  106;  Conservative   56;  Mismatches  146;  Indels  101;  Gaps   17;

Qy         11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASAAE 70
Db        388 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 440

Qy         71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
Db        441 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 493

Qy        121 ---LLLYS----SHQGCSMW-------------MIQHCARDVLEALAFLHHEGYVHADLK 160
Db        494 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 553

Qy        161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
Db        554 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 608

Qy        215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272
Db        609 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 664

Qy        273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH---------- 311
```

```
Db        665 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 720

Qy        312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351
Db        721 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 763


RESULT 13
Q94656
ID   Q94656      PRELIMINARY;      PRT;   826 AA.
AC   Q94656;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mitogen-activated protein kinase 1, serine/threonine protein
DE   kinase.
GN   Name=pfmap1;
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IT04;
RX   MEDLINE=97080491; PubMed=8921836;
RA   Doerig C.D., Parzy D., Langsley G., Horrocks P., Carter R.,
RA   Doerig C.D.;
RT   "A map kinase homologue from the human malaria parasite Plasmodium
RT   falciparum.";
RL   Gene 177:1-6(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IT04;
RA   Doerig C.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; X82646; CAA57972.1; -.
DR   PIR; JC5153; JC5153.
DR   HSSP; Q16539; 1KV1.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR003527; MAP_kin.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS01351; MAPK; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   826 AA;  96645 MW;  A668EF77E734A38A CRC64;

  Query Match           11.0%;  Score 243.5;  DB 2;  Length 826;
  Best Local Similarity   25.8%;  Pred. No. 8.9e-10;
  Matches  105;  Conservative   58;  Mismatches  145;  Indels   99;  Gaps   20;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASA--AEYGFRKERAAL 80
Db         23 YDILKKVGKGAYGVVFKGRCKKNKNIV--AVKKIF-----GAFQNCTDAQRTFR-EIIFL 74

Qy         81 EQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSEL----LLYSSHQGCSMWMIQ 136
Db         75 YELNGHDNIIKLMDVIK---AKNDNDIYLIFDFMETDLHEVIKADLLEEIHK-------K 124

Qy        137 HCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEGNQDV 189
Db        125 YIIYQLLRALKYIHSGGLLHRDIKPSNIL--VNSECHIKVADFGLARSISTHVNENKVPI 182
```

```
Qy        190 --KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL---KHT 244
              |: |  |||||           |   |  |  ||:|||| |: |:  |  |     |
Db        183 LTDYVATRWYRAPEI----------LLGSTHYTEDVDMWSLGCIMGELLCGKPLFTGNST 232

Qy        245 VRSQE------WKANSSAIID--HIFASKAVVN-AAIPAYHLR-----------DLIKSM 284
              :   |       | |   | |    || | : :   :   :|:           ||:: :
Db        233 MNQLEKIIQVIGKPNKKDIEDIRSPFAEKIISSFVDLKKKNLKDICYKASNESLDLLEKL 292

Qy        285 LHDDPSRRIPAEMAL----CSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEY 340
              |  :||:|| || ||       | ||   |    :: :|
Db        293 LQFNPSKRISAENALKHKYVEEFHSIIDEPTCRHIITIP---------------------- 331

Qy        341 EDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQK 387
                : :: |     |  ||  :: : |     |     | |:||  :|
Db        332 --INDNTKYRVNFYRNVVYFVIMRRNKFHSNVL----NQGESKKEEK 372


RESULT 14
Q8ILF0
ID   Q8ILF0      PRELIMINARY;      PRT;   914 AA.
AC   Q8ILF0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mitogen-activated protein kinase 1.
GN   ORFNames=PF14_0294;
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22255705; PubMed=12368864;
RA   Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA   Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA   Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA   Chan M.S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA   Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA   Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA   McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA   Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA   Fraser C.M., Barrell B.;
RT   "Genome sequence of the human malaria parasite Plasmodium
RT   falciparum.";
RL   Nature 419:498-511(2002).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AE014820; AAN36907.1; -.
DR   HSSP; Q16539; 1KV1.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR003527; MAP_kin.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS01351; MAPK; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   914 AA;  107277 MW;  CA4F486AB051F472 CRC64;

  Query Match           11.0%;  Score 243.5;  DB 2;  Length 914;
  Best Local Similarity  25.8%;  Pred. No. 1e-09;
  Matches  105;  Conservative   58;  Mismatches  145;  Indels   99;  Gaps   20;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASA--AEYGFRKERAAL 80
              : :  ::| |:   |:: ||  |      |:|:      ||  :   |:  || |   |
Db         23 YDILKKVGKGAYGVVFKGRCKKNKNIV--AVKKIF-----GAFQNCTDAQRTFR-EIIFL 74

Qy         81 EQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSEL----LLYSSHQGCSMWMIQ 136
              :| || ||: |  |     : |      |: : ::  : |:    ||    |:      :
Db         75 YELNGHDNIIKLMDVIK---AKNDNDIYLIFDFMETDLHEVIKADLLEEIHK-------K 124

Qy        137 HCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEGNQDV 189
              :    :| || ::|  | :| |:|| |||    :||  |: ||||:       |    :
Db        125 YIIYQLLRALKYIHSGGLLHRDIKPSNIL--VNSECHIKVADFGLARSISTHVNENKVPI 182

Qy        190 --KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL---KHT 244
                |: |  |||||           |   |  |  ||:|||| |: |:  |  |     |
Db        183 LTDYVATRWYRAPEI----------LLGSTHYTEDVDMWSLGCIMGELLCGKPLFTGNST 232

Qy        245 VRSQE------WKANSSAIID--HIFASKAVVN-AAIPAYHLR-----------DLIKSM 284
              :   |       | |   | |    || | : :   :   :|:           ||:: :
Db        233 MNQLEKIIQVIGKPNKKDIEDIRSPFAEKIISSFVDLKKKNLKDICYKASNESLDLLEKL 292

Qy        285 LHDDPSRRIPAEMAL----CSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEY 340
              |  :||:|| || ||       | ||   |    :: :|
Db        293 LQFNPSKRISAENALKHKYVEEFHSIIDEPTCRHIITIP---------------------- 331

Qy        341 EDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQK 387
                : :: |     |  ||  :: : |     |     | |:||  :|
Db        332 --INDNTKYRVNFYRNVVYFVIMRRNKFHSNVL----NQGESKKEEK 372


RESULT 15
Q8AYR3
ID   Q8AYR3      PRELIMINARY;      PRT;   395 AA.
AC   Q8AYR3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Calcium/calmodulin-dependent protein kinase (CaM-KI protein).
GN   Name=CaM-KI;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Saneyoshi T., Kume S., Mikoshiba K.;
RT   "Calcium/calmodulin-dependent protein kinase I in Xenopus laevis.";
RL   Comp. Biochem. Physiol. B, Comp. Biochem. 134:499-507(2003).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocytes;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocytes;
RX   MEDLINE=22341132; PubMed=12454917;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
```

```
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocytes;
RA   Klein S., Strausberg R.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AB082999; BAC19847.1; -.
DR   EMBL; BC070745; AAH70745.1; -.
DR   HSSP; Q63450; 1A06.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   395 AA;  43801 MW;  5381D04890B5D3AE CRC64;

  Query Match            11.0%;  Score 242.5;  DB 2;  Length 395;
  Best Local Similarity   27.7%;  Pred. No. 4.2e-10;
  Matches   91;  Conservative   47;  Mismatches  108;  Indels   83;  Gaps   16;

Qy          4 SGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASV-----------YRVRCCGNPGSPPGA 52
              ||  |  :    :|   |::| :  ||:|: : |           : |:|      |  |
Db          9 SGTCWKKQ----VEDIKRMFQFKEVLGTGAFSEVVLAEEKATGKLFAVKCI-----PKKA 59

Qy         53 LKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLL- 111
              ||             |     | | | ::: | ||| |  ::        ||   |:
Db         60 LK-------------GKESSIENEIAVLRKIK-HENIVALEDIY------ESPSHLYLVM 99

Qy        112 ------ELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNIL 165
                   || |  : |   |:     ::       | ||:|:::||  | || |||| |:|
Db        100 QLVSGGELFD-RIVEKGFYTEKDASTL------IRQVLDAVSYLHRLGIVHRDLKPENLL 152

Qy        166 WSAENECFKLI--DFGLSFKEGNQDVKYIQ--TDGYRAPEAELQNCLAQAGLQSDTECTS 221
              : :: |  |::  |||||  ||  ||     .| || |||      |||          :
Db        153 YFSQVEESKIMISDFGLSKMEGKGDVMSTACGTPGYVAPEV-----LAQ------KPYSK 201

Qy        222 AVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAY-----H 276
              ||| ||:|:|   : |           :  | | : : |   ||      | :
Db        202 AVDCWSIGVIAYILLCGY-------PPFYDENDSKLFEQIL--KADYEFDSPYWDDISDS 252

Qy        277 LRDLIKSMLHDDPSRRIPAEMALCSPFFS 305
               :| |:::  ||::|   | ||  |: :
Db        253 AKDFIQNLMEKDPNKRYACEQALRHPWIA 281


Search completed: November 29, 2004, 11:46:52
Job time : 126.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 109.5 Seconds
(without alignments)
1372.672 Million cell updates/sec

Title: US-10-798-532-2
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAFG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2198	99.6	419	4	AAG65767	Aag65767 Human pro
2	2198	99.6	419	5	ABP64797	Abp64797 Human pro
3	2191	99.3	419	4	AAU03534	Aau03534 Human pro
4	2191	99.3	419	4	AAE11769	Aae11769 Human kin
5	2190.5	99.3	418	3	AAY70305	Aay70305 Human wil
6	2187.5	99.2	418	3	AAY70306	Aay70306 Human tra
7	2186	99.1	419	7	ADD48093	Add48093 Rat Prote
8	483.5	21.9	137	5	ABP64662	Abp64662 Human ORF
9	243.5	11.0	638	3	AAG38502	Aag38502 Arabidops
10	243.5	11.0	720	3	AAG38501	Aag38501 Arabidops
11	243.5	11.0	777	3	AAG38500	Aag38500 Arabidops
12	239.5	10.9	544	2	AAW85028	Aaw85028 CDK2-gree
13	238	10.8	298	7	ADD21398	Add21398 Golden ha
14	236	10.7	435	6	AAE31916	Aae31916 Yeast Hog
15	235.5	10.7	367	3	AAB22137	Aab22137 Rat stres
16	235	10.7	435	6	AAE31917	Aae31917 Yeast Hog
17	235	10.7	435	6	AAE31922	Aae31922 Yeast Hog
18	235	10.7	435	6	AAE31921	Aae31921 Yeast Hog
19	235	10.7	435	6	AAE31914	Aae31914 Yeast Hog
20	235	10.7	435	6	AAE31913	Aae31913 Yeast Hog
21	235	10.7	544	2	AAW85029	Aaw85029 Green flu
22	234.5	10.6	433	3	AAB19094	Aab19094 Protein e
23	234.5	10.6	433	8	ADP87472	Adp87472 S cerevis
24	234	10.6	298	2	AAW95689	Aaw95689 Homo sapi
25	234	10.6	298	7	ADD21402	Add21402 Chinese h
26	234	10.6	298	7	ADD21406	Add21406 Rat cycli
27	234	10.6	298	7	ADD21394	Add21394 Mouse cyc
28	234	10.6	416	2	AAW06795	Aaw06795 High Osmo
29	234	10.6	435	6	AAE31910	Aae31910 Yeast Hog
30	234	10.6	435	6	AAE31915	Aae31915 Yeast Hog
31	234	10.6	435	6	AAE31918	Aae31918 Yeast Hog
32	234	10.6	435	6	AAE31919	Aae31919 Yeast Hog
33	234	10.6	435	6	AAE31926	Aae31926 Yeast Hog
34	234	10.6	435	6	AAE31911	Aae31911 Yeast Hog
35	234	10.6	435	6	AAE31909	Aae31909 Yeast Hog
36	234	10.6	435	6	AAE31912	Aae31912 Yeast Hog
37	234	10.6	435	6	AAE31927	Aae31927 Yeast Hog
38	233	10.6	298	6	ABR58686	Abr58686 Human can
39	233	10.6	298	7	ADD21410	Add21410 Goldfish
40	233	10.6	298	7	ADD21392	Add21392 Human cyc
41	233	10.6	298	7	ADF45038	Adf45038 Human kin
42	233	10.6	298	8	ADI57232	Adi57232 NKIAMRE d
43	233	10.6	298	8	ADJ66599	Adj66599 cdk2 prot
44	233	10.6	392	4	ABB63118	Abb63118 Drosophil
45	232.5	10.5	355	4	AAB50055	Aab50055 Murine De

ALIGNMENTS

RESULT 1
AAG65767
ID AAG65767 standard; protein; 419 AA.
XX
AC AAG65767;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human protein kinase polypeptide 13887.
XX
KW Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KW cytostatic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WO200173050-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-US009483.
XX
PR 24-MAR-2000; 2000US-0191846P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers R;
XX
DR WPI; 2001-611632/70.
DR N-PSDB; AAI66828, AAI66829.
XX
PT New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT useful in diagnosis of cancer or cellular proliferation or
PT differentiation disorders and to screen for polypeptide modulators useful
PT to treat such conditions.
XX
PS Claim 4; Fig 10A-C; 169pp; English.
XX
CC The invention provides novel human protein kinase polypeptides, 3714,
CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC protein kinase polypeptides can be expressed by standard recombinant
CC methodology.3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC are useful for diagnostic and screening methods to identify subjects (at
CC risk of) having cancer or cellular proliferation and/or differentiation
CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC modulators are useful for the treatment of cancer, particularly colon
CC cancer or cellular proliferation and/or differentiation disorders. Other
CC disorders associated with 3714, 16742, 23546 or 13887 expression or

CC activity that can be treated include bone related disorders, inflammatory
CC disorders, autoimmune diseases, cardiovascular disorders and liver
CC diseases. The present sequence represents a human protein kinase
CC polypeptide 13887
XX
SQ Sequence 419 AA;

Query Match 99.6%; Score 2198; DB 4; Length 419;
Best Local Similarity 99.8%; Pred. No. 2.1e-229;
Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEEEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
```

RESULT 2
ABP64797
ID ABP64797 standard; protein; 419 AA.
XX
AC ABP64797;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human protein SEQ ID 457.
XX
KW Human; expressed sequence tag; EST; haematopoietic disorder;
KW central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective.
XX
OS Homo sapiens.
XX
PN WO200259260-A2.
XX
PD 01-AUG-2002.
XX
PF 16-NOV-2001; 2001WO-US042950.
XX
PR 17-NOV-2000; 2000US-00714936.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX
DR WPI; 2002-590824/63.
DR N-PSDB; ABQ99383.
XX
PT New isolated polynucleotide, useful in research, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
PT aberrant protein expression or biological activity.
XX
PS Claim 20; SEQ ID NO 457; 394pp; English.
XX
CC The present invention relates to novel human coding sequences (ABQ99268-
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC therapeutic, diagnostic and research methods. The polynucleotides may be
CC used in the field of molecular biology as hybridisation probes, primers
CC for PCR, for chromosome and gene mapping, for the recombinant production
CC of protein, or in generation of anti-sense DNA or RNA. The
CC polynucleotides are useful in diagnostics as expressed sequence tags
CC (ESTs) for identifying expressed genes or for physical mapping of the
CC human genome. The proteins may be used as molecular weight markers, or as
CC nutritional sources or supplements. The proteins may be used to maintain
CC and expand cell population in a totipotential or pluripotential state
CC useful for re-engineering damaged or diseased tissues, transplantation,
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC polynucleotides and proteins are useful for preventing, treating or
CC ameliorating disorders involving aberrant protein expression or
CC biological activity, e.g. haematopoietic disorders, central/peripheral
CC nervous system diseases, mechanical and traumatic disorders, non-healing
CC wounds, immune deficiencies and disorders, infectious diseases caused by
CC viral, bacterial or fungal infection, autoimmune disorders, allergic
CC reactions and conditions, coagulation disorders, or cancer. The
CC polynucleotide sequences of the invention were assembled from ESTs
CC isolated mainly by sequencing by hybridisation, and in some cases,
CC sequences obtained from one or more public databases. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 419 AA;

Query Match 99.6%; Score 2198; DB 5; Length 419;
Best Local Similarity 99.8%; Pred. No. 2.1e-229;
Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEEEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
```

RESULT 3

```
AAU03534
ID   AAU03534 standard; protein; 419 AA.
XX
AC   AAU03534;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human protein kinase #34.
XX
KW   Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW   metabolic disorder; immune related disease; neurological disorder;
KW   neurodegenerative disorder; inflammatory disorder; infectious disease;
KW   reproductive disorder.
XX
OS   Homo sapiens.
XX
PN   WO200138503-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US032085.
XX
PR   24-NOV-1999;   99US-0167482P.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman GD,  Whyte D,  Manning G,  Sudarsanam S,  Martinez R;
PI   Flanagan P,  Clary D;
XX
DR   WPI; 2001-343950/36.
DR   N-PSDB; AAS06734.
XX
PT   Nucleic acids encoding human kinase polypeptides, useful for preventing
PT   diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT   neuronal-associated diseases, and microbial infections.
XX
PS   Claim 7; Fig 2; 433pp; English.
XX
CC   AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel
CC   protein kinases have been identified as members of the tyrosine or
CC   serine/threonine kinase (PTK and STK) families. The polynucleotides
CC   encoding protein kinases and the polypeptides may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate kinase expression. For example, they may be used to treat
CC   cancers (especially cancers of haematopoietic origin), cardiovascular
CC   disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC   immune related diseases (e.g. rheumatoid arthritis), neurological
CC   disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC   Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC   disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC   Additionally, polynucleotides encoding protein kinases may be used for
CC   gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC   polypeptides may be used as antigens in the production of antibodies
CC   against the protein kinases and in assays to identify modulators of
CC   protein kinase expression and activity
XX
SQ   Sequence 419 AA;

  Query Match             99.3%;  Score 2191;  DB 4;  Length 419;
  Best Local Similarity   99.5%;  Pred. No. 1.2e-228;
  Matches  417;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEEEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKGNPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 4
AAE11769
ID   AAE11769 standard; protein; 419 AA.
XX
AC   AAE11769;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human kinase (PKIN)-3 protein.
XX
KW   Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
KW   cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
KW   AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
KW   atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
KW   myasthenia gravis; cirrhosis; cataract; growth and development disorder;
KW   seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW   lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW   asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
KW   antimicrobial; cardiovascular; antiinflammatory; vaccine.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          23. .304
FT                   /note= "Eukaryotic protein kinase domain"
FT   Domain          48. .289
FT                   /note= "Protein kinase domain"
FT   Domain          74. .293
FT                   /note= "Protein kinase domain"
FT   Domain          108. .241
FT                   /note= "Protein kinase domain"
FT   Domain          148. .166
FT                   /note= "Tyrosine kinase catalytic domain"
FT   Domain          223. .245
FT                   /note= "Tyrosine kinase catalytic domain"
FT   Domain          223. .241
FT                   /label= Transmembrane_domain
FT   Domain          273. .295
FT                   /note= "Tyrosine kinase catalytic domain"
FT   Region          305. .344
FT                   /note= "Serine/threonine protein kinase"
FT   Region          319. .404
FT                   /note= "Ribonucleoprotein repeat"
FT   Region          320. .412
FT                   /note= "Splicing factor like protein"
XX
PN   WO200181555-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-US012992.
XX
PR   20-APR-2000; 2000US-0199021P.
PR   28-APR-2000; 2000US-0200226P.
PR   05-MAY-2000; 2000US-0202339P.
```

```
PR   11-MAY-2000; 2000US-0203505P.
PR   18-MAY-2000; 2000US-0205564P.
PR   26-MAY-2000; 2000US-0207739P.
PR   01-JUN-2000; 2000US-0208795P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
PI   Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Walia NK, Yao MG;
PI   Patterson C, Burrill JD, Marcus GA, Zingler KA, Recipon SA, Lu Y;
PI   Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
PI   Walsh RT, Ramkumar J, Borowsky ML, Au-Young J, Hillman JL;
PI   Gururajan R;
XX
DR   WPI; 2001-611740/70.
DR   N-PSDB; AAD18818.
XX
PT   Human kinases and nucleic acids, useful for preventing diagnosing and
PT   treating cancers, inflammation and immune disorders.
XX
PS   Claim 1; Page 125-126; 166pp; English.
XX
CC   The present invention relates to human kinases (PKIN) and the nucleic
CC   acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
CC   used in the prevention, diagnosis and treatment of diseases cancers,
CC   adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
CC   acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
CC   gout, microbial infections, cardiovascular disease and/or inflammation,
CC   myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
CC   infarction, cataract, growth and development disorder, seizure disorder,
CC   pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
CC   disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
CC   PKIN may be used to treat disorders associated with decreased PKIN
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of PKIN by expressing inactive proteins or to
CC   supplement the patients own production of PKIN. PKIN nucleic acids may be
CC   used to produce the PKIN polypeptide, by inserting the nucleic acids into
CC   a host cell and culturing the cell to express the protein. PKIN nucleic
CC   acid and its complementary sequences may also be used as DNA probes in
CC   diagnostic assays to detect and quantitate the presence of similar
CC   nucleic acid sequences in samples and therefore which patients may be in
CC   need of restorative therapy. The present sequence is human PKIN-3 protein
XX
SQ   Sequence 419 AA;

  Query Match             99.3%;  Score 2191;  DB 4;  Length 419;
  Best Local Similarity   99.5%;  Pred. No. 1.2e-228;
  Matches  417;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEEEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKGNPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 5
AAY70305
ID   AAY70305 standard; protein; 418 AA.
XX
AC   AAY70305;
XX
DT   06-JUN-2000  (first entry)
XX
DE   Human wild type serine/threonine kinase KIS (hKIS) protein.
XX
KW   KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
KW   cyclin-dependent kinase inhibitor; CKI; p27; modulator; treatment;
KW   cell proliferative disease; vascular disorder; gene therapy; restenosis;
KW   atherosclerosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 186
FT                   /note= "Encoded by AATCAG"
XX
PN   WO200011165-A1.
XX
PD   02-MAR-2000.
XX
PF   20-AUG-1999;   99WO-US018903.
XX
PR   21-AUG-1998;   98US-0097710P.
XX
PA   (NABE/) NABEL G J.
PA   (NABE/) NABEL E G.
XX
PI   Nabel GJ,  Nabel EG;
XX
DR   WPI; 2000-237648/20.
DR   N-PSDB; AAZ51355.
XX
PT   Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT   for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT   proliferation.
XX
PS   Example 1; Page 57-58; 70pp; English.
XX
CC   The present protein sequence is the wild type human KIS (hKIS), isolated
CC   from a human B-cell library, using a yeast two hybrid screening system.
CC   hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
CC   cyclin-dependent kinase inhibitor (CKI), p27. hKIS controls cell
CC   proliferation and is localised predominantly in the nucleus. It binds to
CC   CKI p27 and inhibits its ability to arrest cells in G1 phase. It has an
CC   amino terminal serine/threonine kinase consensus region and a carboxy
CC   terminus with 42% sequence similarity to hU2AF65, a subunit of the
CC   splicing factor U2AF. The hKIS sequences are used to modulate cell
CC   proliferation and treat cell proliferative and vascular diseases. The DNA
CC   sequence may be used in gene therapy to treat vascular disorders such as
CC   restenosis or atherosclerosis
XX
SQ   Sequence 418 AA;

  Query Match             99.3%;  Score 2190.5;  DB 3;  Length 418;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-228;
  Matches  418;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGN-DVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 239

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 299

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 359

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 418
```

RESULT 6
AAY70306
ID AAY70306 standard; protein; 418 AA.
XX
AC AAY70306;
XX
DT 06-JUN-2000 (first entry)
XX
DE Human transdominant mutant serine/threonine kinase KIS (hKIS) protein.
XX
KW KIS; human; hKIS; serine/threonine kinase; transdominant mutant; p27;
KW G1 phase; cell proliferation; cyclin-dependent kinase inhibitor; CKI;
KW modulator; treatment; cell proliferative disease; vascular disorder;
KW gene therapy; atherosclerosis; restenosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 54
FT /note= "Wild type Lys substituted with Arg"
FT Misc-difference 186
FT /note= "Encoded by AATCAG"
XX
PN WO200011165-A1.
XX
PD 02-MAR-2000.
XX
PF 20-AUG-1999; 99WO-US018903.
XX
PR 21-AUG-1998; 98US-0097710P.
XX
PA (NABE/) NABEL G J.
PA (NABE/) NABEL E G.
XX
PI Nabel GJ, Nabel EG;
XX
DR WPI; 2000-237648/20.
DR N-PSDB; AAZ51356.
XX
PT Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT proliferation.
XX
PS Example 1; Page 59-61; 70pp; English.
XX
CC The present protein sequence is the transdominant mutant human KIS
CC (hKIS), constructed by site directed mutagenesis. A single nucleotide
CC substitution (A to G) results in a lysine to arginine change in the
CC protein sequence. hKIS is a serine/threonine kinase, that acts as an
CC inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
CC controls cell proliferation and is localised predominantly in the
CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC G1 phase. The hKIS sequences are used to modulate cell proliferation and
CC treat cell proliferative and vascular diseases. The polynucleotide
CC sequence may be used in gene therapy to treat vascular disorders such as
CC restenosis or atherosclerosis
XX
SQ Sequence 418 AA;

Query Match 99.2%; Score 2187.5; DB 3; Length 418;
Best Local Similarity 99.5%; Pred. No. 2.9e-228;
Matches 417; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```
Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGN-DVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 239

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 299

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 359

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 418
```

RESULT 7
ADD48093
ID ADD48093 standard; protein; 419 AA.
XX
AC ADD48093;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein AAC53031, SEQ ID NO 13791.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.

```
DR   GENBANK; AAC53031.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a rat protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 419 AA;

  Query Match             99.1%;  Score 2186;  DB 7;  Length 419;
  Best Local Similarity   99.0%;  Pred. No. 4.3e-228;
  Matches  415;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGTPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||| ||:||||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 8
ABP64662
ID   ABP64662 standard; protein; 137 AA.
XX
AC   ABP64662;
XX
DT   04-NOV-2002  (first entry)
XX
DE   Human ORF1032.
XX
KW   Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW   Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW   human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW   cancer; cardiovascular disease; allergy; autoimmune disease;
KW   wound healing; blood coagulation disorder; inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   US2002082206-A1.
XX
PD   27-JUN-2002.
XX
PF   30-MAY-2001; 2001US-00867550.
XX
PR   30-MAY-2000; 2000US-0208427P.
XX
PA   (LEAC/) LEACH M D.
PA   (MEHR/) MEHRABAN F.
PA   (CONL/) CONLEY P B.
PA   (TOPP/) TOPPER J N.
PA   (LAWD/) LAW D.
XX
PI   Leach MD,  Mehraban F,  Conley PB,  Topper JN,  Law D;
XX
DR   WPI; 2002-626554/67.
DR   N-PSDB; ABQ99225.
XX
PT   New polypeptide designated ORFX are present in human atherogenic cells
PT   and are useful to prevent and treat ORFX-associated disorders including
PT   cancer, allergy, wound healing or autoimmune, cardiovascular or
PT   inflammatory disease.
XX
PS   Claim 10; SEQ ID NO 2064; 78pp; English.
XX
CC   The present invention relates to novel human ORFX polypeptides and their
CC   coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC   were discovered in human atherogenic cells, in particular in platelets
CC   and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC   many other tissues as well. Atherogenic cells are cells which have the
CC   potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC   nucleic acids are useful for treating or preventing a pathological
CC   condition associated with an ORFX-associated disorder, e.g. cancer,
CC   cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC   coagulation disorders or inflammatory disorders. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/sequence.html?DocID=20020082206
XX
SQ   Sequence 137 AA;

  Query Match             21.9%;  Score 483.5;  DB 5;  Length 137;
  Best Local Similarity   80.0%;  Pred. No. 1e-43;
  Matches   96;  Conservative    1;  Mismatches   12;  Indels   11;  Gaps    1;

Qy         63 GAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELL 122
              |  |:|   | |   |           |||||||||||||||||||||||||||||||
Db         29 GRGAAAFSGGLRAAVA----------VTLYGVFTIHFSPNVPSRCLLLELLDVSVSELL 77

Qy        123 LYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSF 182
              |||||||||||||||||||||||||||||||||||| ||||||||||  ||||||||
Db         78 LYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLXPRNILWSAENECXTLIDFGLSF 137


RESULT 9
AAG38502
ID   AAG38502 standard; protein; 638 AA.
```

XX
AC AAG38502;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47508.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.

```
PR   25-AUG-1999;   99US-0150566P.
PR   26-AUG-1999;   99US-0150884P.
PR   27-AUG-1999;   99US-0151065P.
PR   27-AUG-1999;   99US-0151066P.
PR   27-AUG-1999;   99US-0151080P.
PR   30-AUG-1999;   99US-0151303P.
PR   31-AUG-1999;   99US-0151438P.
PR   01-SEP-1999;   99US-0151930P.
PR   07-SEP-1999;   99US-0152363P.
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match           11.0%;  Score 243.5;  DB 3;  Length 638;
  Best Local Similarity  25.9%;  Pred. No. 1.4e-16;
  Matches  106;  Conservative   56;  Mismatches  146;  Indels  101;  Gaps   17;

Qy         11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASAAE 70
              :|  | :  |:|:|   ::||| |: |::|    :      |||:    |   | |
Db        249 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 301

Qy         71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
              ||| :|   |::|:|  ||: |           |  : || |:|: ::|
Db        302 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 354

Qy        121 ---LLLYS----SHQGCSMW------------MIQHCARDVLEALAFLHHEGYVHADLK 160
              :| |      :|     |             ::   : :|:|:  :| |  ||:|||
Db        355 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 414

Qy        161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
              | | |        ||||||:: |  | |   || |       | :||| : |   : |
Db        415 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 469

Qy        215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272
              :  :|    |:|||| || :|  | :         | |    | |  |      : |
Db        470 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 525

Qy        273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH---------------------- 311
                | ||:|  |  | ::|       |  || : |  ||
Db        526 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 581

Qy        312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351
                      :| |:  | |: |      :| | : :: : ::  | | |
Db        582 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 624


RESULT 10
AAG38501
ID   AAG38501 standard; protein; 720 AA.
XX
AC   AAG38501;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 47507.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
```

PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.0%; Score 243.5; DB 3; Length 720;
Best Local Similarity 25.9%; Pred. No. 1.7e-16;

```
  Matches  106;  Conservative   56;  Mismatches  146;  Indels  101;  Gaps   17;

Qy          11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASAAE 70
               :|  | :  |:|:|   ::||| |: |::|    :       |||:    |   | |
Db         331 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 383

Qy          71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
               ||| :|   |::|:|  ||: |           |  : || |:|: ::|
Db         384 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 436

Qy         121 ---LLLYS----SHQGCSMW------------MIQHCARDVLEALAFLHHEGYVHADLK 160
                  :| |     :|     |              ::   : :|:|:  :| |  ||:|||
Db         437 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 496

Qy         161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
               | | |        ||||||:: |  | |   || |       | :||| : |   : |
Db         497 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 551

Qy         215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272
               :  :|    |:|||| || :|  | :         | |    | |  |     : |
Db         552 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 607

Qy         273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH---------------------- 311
                   | ||:|  |  | ::|      |  || : |  ||
Db         608 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 663

Qy         312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351
                       :| |:  | |: |      :| | : :: : ::  | | |
Db         664 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 706


RESULT 11
AAG38500
ID   AAG38500 standard; protein; 777 AA.
XX
AC   AAG38500;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 47506.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
```

PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.0%; Score 243.5; DB 3; Length 777;
Best Local Similarity 25.9%; Pred. No. 1.9e-16;
Matches 106; Conservative 56; Mismatches 146; Indels 101; Gaps 17;

```
Qy         11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPGTTGAAASAAE 70
              :|  | :  |:|:|   ::||| |: |::|    :       |||:   |   | |
Db        388 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 440

Qy         71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
              ||| :|   |::|:|  ||: |           |  : || |:|: ::|
Db        441 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 493

Qy        121 ---LLLYS----SHQGCSMW------------MIQHCARDVLEALAFLHHEGYVHADLK 160
                 :| |     :|     |              ::   : :|:|:  :| |  ||:|||
Db        494 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 553

Qy        161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
              | | |        ||||||:: |  | |   || |       | :||| : |    : |
Db        554 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 608

Qy        215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272
              :  :|    |:|||| || :|  | :         | |    | |  |      : |
Db        609 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 664

Qy        273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH---------------------- 311
                 | ||:|  |  | ::|      |  || : |  ||
Db        665 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 720

Qy        312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351
                      :| |:  | |: |      :| | : :: : ::  | | |
Db        721 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 763
```

RESULT 12
AAW85028
ID AAW85028 standard; protein; 544 AA.
XX
AC AAW85028;
XX
DT 17-OCT-2003 (revised)
DT 08-FEB-1999 (first entry)
XX
DE CDK2-green fluorescent protein usion product.
XX
KW Human; CDK2 gene; fusion protein; green fluorescent protein; GFP;
KW intracellular signalling; chimera.
XX
OS Aequorea victoria.
OS Homo sapiens.
OS Chimeric.
XX
PN WO9845704-A2.
XX
PD 15-OCT-1998.
XX
PF 07-APR-1998; 98WO-DK000145.
XX
PR 07-APR-1997; 97DK-00000392.

```
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Thastrup O,  Petersen Bjorn S,  Tullin S,  Kasper A,  Scudder K;
XX
DR   WPI; 1998-594491/50.
DR   N-PSDB; AAV71073.
XX
PT   Determining effect on signalling pathways in live cells from
PT   redistribution of luminophores - specifically fusions of green
PT   fluorescent protein with a signalling component, and new apparatus,
PT   particularly for identifying toxins and potential therapeutic agents.
XX
PS   Example 12; Page 208-209; 326pp; English.
XX
CC   The present sequence represents a human CDK2-green fluorescent protein
CC   (GFP) fusion protein. The fusion protein is used in an assay to exemplify
CC   the invention. The specification describes how quantitative information
CC   about the influence of a molecule on a cellular response is obtained by
CC   recording the variation, caused by the molecule, on mechanically intact
CC   living cells, in the spatially distributed light emitted from a
CC   luminophore present in the cells. The variation in light emission is
CC   processed to provide information that correlates spatial distribution to
CC   the degree of the molecule. The method is used to identify agents that
CC   (in)directly affect intracellular signalling, especially to screen for
CC   potential therapeutic agents or toxins, and to identify new drug targets.
CC   (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 544 AA;

  Query Match             10.9%;  Score 239.5;  DB 2;  Length 544;
  Best Local Similarity   27.0%;  Pred. No. 3e-16;
  Matches  112;  Conservative   52;  Mismatches  154;  Indels   97;  Gaps   18;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQF-LPPGTTGAAASAAEYGFRKERAALE 81
              :|   ::| |:   ||: |           |||:  |   | |  ::|      :| : |:
Db          4 FQKVEKIGEGTYGVVYKAR--NKLTGEVVALKKIRLDTETEGVPSTAI-----REISLLK 56

Qy         82 QLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARD 141
              :|  | ||| |  |  ||    :   |: | |   : : :  |:  |  : :|:
Db         57 EL-NHPNIVKLLDV--IHTENKL---YLVFEFLHQDLKKFMDASALTGIPLPLIKSYLFQ 110

Qy        142 VLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY---IQTDGYR 198
              :|: ||| |    :| |||||:|:| : |    || ||||:   |      |   : |  ||
Db        111 LLQGLAFCHSHRVLHRDLKPQNLLINTEG-AIKLADFGLARAFGVPVRTYTHEVVTLWYR 169

Qy        199 APEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAII 258
              |||           |      ::|||:|||| |  ||          |   :  :|   |
Db        170 APEI----------LLGSKYYSTAVDIWSLGCIFAEM--------VTRRALFPGDSE--I 209

Qy        259 DHIFASKAVVN----------AAIPAYHL--------------------RDLIKSMLHD 287
              | :|    :            ::| |                      | |:  |||
Db        210 DQLFRIFRTLGTPDEVVWPGVTSMPDYKPSFPKWARQDFSKVVPPLDEDGRSLLSQMLHY 269

Qy        288 DPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDV 347
              ||::|| |: ||  |||        :| |  | | ||| :                |   |
Db        270 DPNKRISAKAALAHPFF--------QD-VTKPVPHLRLWD-------------PPVATMV 307

Qy        348 KEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVAT 402
               :  : :  || :||  :    |  |   :  |:  |    ||      ||:  |
Db        308 SKGEELFTGVVPILVELDGDVNGHKF-SVSGEGEGDATYGKLT-----LKFICTT 356


RESULT 13
ADD21398
ID   ADD21398 standard; protein; 298 AA.
XX
AC   ADD21398;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Golden hamster cyclin dependent kinase (cdk) 2 protein.
XX
KW   continual growth; cultured cell; cyclin dependent kinase; cdk4; cdk2;
KW   cdk6; activating mutation; cell growth; cell division; cell cycle;
KW   cancer-causing agent; continual growth-induced cell; enzyme;
KW   golden hamster.
XX
OS   Mesocricetus auratus.
XX
PN   WO2003044169-A2.
XX
PD   30-MAY-2003.
XX
PF   15-NOV-2002; 2002WO-US036729.
XX
PR   15-NOV-2001; 2001US-0334760P.
XX
PA   (UTEM ) UNIV TEMPLE.
XX
PI   Reddy PE,  Rane SG,  Mettus RV;
XX
DR   WPI; 2003-449813/42.
DR   N-PSDB; ADD21399.
XX
PT   A composition for reversibly inducing continual growth in normal cells
PT   comprises a cyclin dependent kinase protein (e.g. cdk4, cdk2 or cdk6) or
PT   its active fragment, derivative, homolog or analog, having an activating
PT   mutation.
XX
PS   Disclosure; Page 104; 77pp; English.
XX
CC   This invention relates to a novel composition for inducing a reversible
CC   state of a continual growth in cultured cells and comprises at least one
CC   compound comprising a cyclin dependent kinase (cdk)4, cdk2 or cdk6
CC   protein having an activating mutation. Growth and division of living
CC   cells involve a regular series of events and processes that comprise the
CC   cell cycle. Cyclin dependent kinases cdk2, cdk4 and cdk6 are involved in
CC   the control of G1, the point at which cells irrevocably commit to DNA
CC   synthesis and thus enter the cell cycle. The invention is useful in
CC   reversibly inducing continual growth in normal cells and may allow the
CC   screening of cancer-causing agents with the continual growth-induced
CC   cells. The present sequence is that of the golden hamster cdk2 protein
CC   related to the invention. Note: Due to an error in the specification or
CC   sequence listing, the Seq ID numbers given in the disclosure do not
CC   correspond to those given in the sequence listing. It is therefore
CC   unclear which Seq ID number corresponds to which sequence and exactly
CC   which sequence is being claimed.
XX
SQ   Sequence 298 AA;

  Query Match             10.8%;  Score 238;  DB 7;  Length 298;
  Best Local Similarity   29.3%;  Pred. No. 1.7e-16;
  Matches   99;  Conservative   44;  Mismatches  117;  Indels   78;  Gaps   16;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQF-LPPGTTGAAASAAEYGFRKERAALE 81
              :|   ::| |:   ||: :           |||:  |   | |  ::|      :| : |:
Db          4 FQKVEKIGEGTYGVVYKAK--NKLTGEVVALKKIRLDTETEGVPSTAI-----REISLLK 56

Qy         82 QLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARD 141
              :|  | ||| |  |  ||    :    |: |||   : : :  |:  |  : :|:
Db         57 EL-NHPNIVKLLDV--IHTENKL---YLVFELLHQDLKKFMDASAVTGIPLPLIKSYLFQ 110

Qy        142 VLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY---IQTDGYR 198
              :|: ||| |    :| ||||:|:| :||    || ||||:   |      |   : |  ||
Db        111 LLQGLAFCHSHRVLHRDLKPQNLLINAEGS-IKLADFGLARAFGVPVRTYTHEVVTLWYR 169

Qy        199 APEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAII 258
              |||  |  |            ::|||:|||| |  ||          |   :  :|   |
Db        170 APEI-LLGC---------KYYSTAVDIWSLGCIFAEM--------VTRRALFPGDSE--I 209

Qy        259 DHIFASKAVVN----------AAIPAYHL--------------------RDLIKSMLHD 287
              | :|    :            ::| |                      | |:  |||
Db        210 DQLFRIFRTLGTPDEVVWPGVTSMPDYKPSFPKWARQDFSKVVPPLDEDGRSLLSQMLHY 269

Qy        288 DPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRL 325
```

```
                ||::|| |: ||  |||        :| |  | | |||
Db        270 DPNKRISAKAALAHPFF--------QD-VTKPVPHLRL 298


RESULT 14
AAE31916
ID   AAE31916 standard; protein; 435 AA.
XX
AC   AAE31916;
XX
DT   07-MAR-2003  (first entry)
XX
DE   Yeast Hog1 mutant protein (Y68H).
XX
KW   Yeast; Hog1; enzyme; mitogen activated protein kinase; MAPK pathway;
KW   drug design; signal transduction; neoplasia; cancer; inflammation;
KW   degenerative disease; immunological disorder; therapy; mutant; mutein.
XX
OS   Saccharomyces cerevisiae.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 68
FT                   /note= "Wild type Tyr substituted with His"
XX
PN   WO200281746-A2.
XX
PD   17-OCT-2002.
XX
PF   04-APR-2002; 2002WO-IL000281.
XX
PR   05-APR-2001; 2001IL-00142481.
XX
PA   (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI   Engelberg D,  Bell M,  Levitzki A;
XX
DR   WPI; 2003-067522/06.
XX
PT   Screening for constitutively activated mutants of the mitogen activated
PT   protein kinase (MAPK) pathway member, useful for screening drugs for
PT   treating e.g. cancer, by employing a mutant yeast strain devoid of an
PT   upstream kinase.
XX
PS   Claim 22; Page; 123pp; English.
XX
CC   The invention relates to a method of screening for constitutively active
CC   mutant of a desired eukaryotic mitogen activated protein kinase (MAPK)
CC   pathway member (such as Hog1 in yeast). The method involves employing a
CC   mutant yeast strain devoid of an upstream kinase which is capable of
CC   activating the MAPK pathway member of a yeast MAPK pathway that is
CC   equivalent or corresponding to the desired eukaryotic MAPK pathway
CC   member. The method is useful in drug design particularly for screening
CC   constitutively active mutants of MAPKs and substances that are inhibitors
CC   of MAPK signal transduction pathways. MAPK pathway inhibitors are used
CC   for treating pathological conditions e.g. neoplasia, cancer,
CC   inflammation, degenerative diseases or immunological disorders. The
CC   present sequence is yeast Hog1 mutant protein. Note: This sequence is not
CC   shown in the specification but is derived from the yeast Hog1 wild-type
CC   protein shown as SEQ ID NO: 1 (AAE31909) in figure 3A of the
CC   specification
XX
SQ   Sequence 435 AA;

  Query Match             10.7%;  Score 236;  DB 6;  Length 435;
  Best Local Similarity   25.7%;  Pred. No. 5.1e-16;
  Matches   98;  Conservative   53;  Mismatches  140;  Indels   90;  Gaps   18;

Qy         17 EAFGRLWQVQSR------LGSGSSASVYRVRCCGNP--GSPPGALKQFLPPGTTGAAASA 68
              : || :::: :|      :| |:   |    |       | | |:|: : | :|   |
Db         11 QIFGTVFEITNRYNDLNPVGMGAPGLV----CSATDTLTSQPVAIKKIMKPFSTAVLAKR 66

Qy         69 AEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQ 128
                :|   |: |: | |:: |  :|    || :     : ||    :  ||     :
Db         67 T----HRELKLLKHLR-HENLICLQDIF---LSP-LEDIYFVTELQGTDLHRLL---QTR 114

Qy        129 GCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQD 188
               :|:    :|  | ::|  | :| |||| ||| : ||   |: ||||:  :  |
Db        115 PLEKQFVQYFLYQILRGLKYVHSAGVIHRDLKPSNILIN-ENCDLKICDFGLARIQDPQM 173

Qy        189 VKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL----KHT 244
               |: |  |||||  |          : | |    ||:|| | |  ||  |  |      |
Db        174 TGYVSTRYYRAPEIML------TWQKYDVE----VDIWSAGCIFAEMIEGKPLFPGKDH- 222

Qy        245 VRSQEWKANSSAIIDHIFAS--KAVVNAAIPAYHLR------------------------ 278
                   :  :||  :  |  | |:|       |:
Db        223 -------VHQFSIITDLLGSPPKDVINTICSENTLKFVTSLPHRDPIPFSERFKTVEPDA 275

Qy        279 -DLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVM-----------LPTPVLRL 325
               ||:: ||  || :|| |  ||  | :| |:    :: |              ||    |:
Db        276 VDLLEKMLVFDPKKRITAADALAHP-YSAPYHDPTDEPVADAKFDWHFNDADLPVDTWRV 334

Qy        326 L---NVLDDDYLGNEEEYEDV 343
               :   :||   :|  :   |:
Db        335 MMYSEILDFHKIGGSDGQIDI 355


RESULT 15
AAB22137
ID   AAB22137 standard; protein; 367 AA.
XX
AC   AAB22137;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Rat stress-activated protein kinase 3.
XX
KW   Rat; stress-activated protein kiase 3; SAPK3; PDZ domain; GLGF loop;
KW   DHF domain; muscular dystrophy; ERK6; p38gamma.
XX
OS   Rattus sp.
XX
PN   WO200048002-A1.
XX
PD   17-AUG-2000.
XX
PF   09-FEB-2000; 2000WO-GB000374.
XX
PR   09-FEB-1999;   99GB-00002696.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
XX
PI   Spillantini MG,  Goedert M,  Hasegawa M,  Buee-Scherrer V,  Thomas G;
PI   Cohen P,  Cuenda A;
XX
DR   WPI; 2000-558226/51.
DR   N-PSDB; AAA90869.
XX
PT   Novel methods for identifying compounds useful for treating muscular
PT   dystrophy comprising identifying modulators of protein kinase and PDZ
PT   domain binding.
XX
PS   Disclosure; Fig 6; 102pp; English.
XX
CC   The present sequence is the protein sequence of the rat stress-activated
CC   protein kinase 3 (SAPK3, also known as ERK6 and p38gamma) enzyme. It was
CC   used to demonstrate the methods and peptides of the invention, which
CC   involve the identification of compounds which are able to modulate the
CC   interaction between PDZ domains (also known as DHF or GLGF domains) and
CC   protein kinases. The invention also comprises peptide fragments derived
CC   from stress-activated protein kinases which are involved in protein
CC   phosphorylation. The compounds and peptides can be used in the treatment
CC   of muscular diseases, for example muscular dystrophy, or in cases where
CC   modulation of phosphorylation of proteins with PDZ domains or modulation
CC   of signalling via activin receptors or voltage gated channels is
CC   necessary, or in instances where the localisation of SAPK3 at the
```

```
CC   neuromuscular junction or sarcolemma needs to be disrupted
XX
SQ   Sequence 367 AA;

  Query Match             10.7%;  Score 235.5;  DB 3;  Length 367;
  Best Local Similarity   22.8%;  Pred. No. 4.4e-16;
  Matches   96;  Conservative   67;  Mismatches  127;  Indels  131;  Gaps   19;

Qy          10 AEPPRPLEAFGR------LWQVQS------RLGSGSSASVYRVRCCGNPGSPPG---ALK 54
               : ||   : | |       |:|::       :|||:  :|     |    |  |   |:|
Db           2 SSPPPARKGFYRQEVTKTAWEVRAVYQDLQPVGSGAYGAV-----CSAVDSRTGNKVAIK 56

Qy          55 QFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRC------ 108
               :   |  :   |  |     :|   |: :: | |:: |  |||       |
Db          57 KLYRPFQSELFAKRA----YRELRLLKHMR-HENVIGLLDVFT-------PDETLDDFTD 104

Qy         109 --LLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILW 166
                 |::  :   : :|:    |:  |   ||     :|: | ::|  | :| |||| |:
Db         105 FYLVMPFMGTDLGKLM---KHETLSEDRIQFLVYQMLKGLKYIHAAGVIHRDLKPGNL-- 159

Qy         167 SAENEC-FKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDL 225
               :   :|  |::||||: :  ::   |: |  |||||  |          :    |  ||:
Db         160 AVNEDCELKILDFGLARQADSEMTGYVVTRWYRAPEVIL----------NWMRYTQTVDI 209

Qy         226 WSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNA-------------- 270
               ||:| |: || :|  |         :| |     ||:   | ::
Db         210 WSVGCIMAEMITGKIL--------FKGN-----DHLDQLKEIMKVTGTPPPEFVQKLQSA 256

Qy         271 -------AIPAYHLRD-------------LIKSMLHDDPSRRIPAEMALCSPFFSIPFA 309
                    :|    :|               |:: ||  |  :|: |  ||  |:|
Db         257 EAKNYMEGLPELEKKDFASVLTNASPQAVNLLEKMLVLDAEQRVTAAEALAHPYFE---- 312

Qy         310 PHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVK---EECQK--YGPVVSLLVPK 364
                         :: | :     ::|:|  :||    || ::  |  |:|   |:
Db         313 ----------------SLRDTEDEPKAQKYDDSFDDVDRTLEEWKRVTYKEVLSFKPPR 355

Qy         365 E 365
               :
Db         356 Q 356


Search completed: November 29, 2004, 11:30:36
Job time : 112.5 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:30 ; Search time 5310 Seconds
(without alignments)
8646.716 Million cell updates/sec

Title: US-10-798-532-3
Perfect score: 1260
Sequence: 1 atggcgggatccggctgcgc.....tgtatcaaaccttgctttaa 1260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1113.4	88.4	1763	3	AK013347	AK013347 Mus muscu
	2	886.6	70.4	2846	3	AK030152	AK030152 Mus muscu
	3	861.8	68.4	1079	4	BM468107	BM468107 AGENCOURT
	4	723.8	57.4	917	4	BM451184	BM451184 AGENCOURT
	5	622.8	49.4	1015	6	BY712344	BY712344 BY712344
	6	619.2	49.1	657	4	BG547675	BG547675 602575520
	7	617.8	49.0	899	2	BE895119	BE895119 601436041
	8	616.4	48.9	772	4	BI562420	BI562420 603256316
	9	608.8	48.3	932	5	BU541306	BU541306 AGENCOURT
	10	595	47.2	875	4	BI259568	BI259568 602968330
	11	591.6	47.0	834	7	CF593363	CF593363 AGENCOURT
	12	583.8	46.3	906	4	BG386905	BG386905 602454704
	13	583.6	46.3	682	4	BI333581	BI333581 602997069
	14	578	45.9	888	5	BQ225071	BQ225071 AGENCOURT
	15	560	44.4	602	7	CO737403	CO737403 SlLT02c17
	16	557.6	44.3	779	5	BU447831	BU447831 603767571
	17	556.4	44.2	876	5	BU184346	BU184346 AGENCOURT
	18	549.2	43.6	844	4	BG474616	BG474616 602517423
	19	519.8	41.3	824	2	BE568528	BE568528 601342425
	20	516.8	41.0	823	7	CO771206	CO771206 testis_ES
	21	516	41.0	781	7	CN232453	CN232453 WLB074D11
	22	498.6	39.6	717	5	BU385761	BU385761 603582650
	23	497.8	39.5	670	7	CN359673	CN359673 170004245
c	24	496.8	39.4	523	2	BF929410	BF929410 IL2-NT020
	25	495.2	39.3	896	7	CO773091	CO773091 testis_ES
	26	494.2	39.2	1003	5	BQ922020	BQ922020 AGENCOURT
	27	493	39.1	928	5	BQ424438	BQ424438 AGENCOURT
	28	482.4	38.3	716	7	CF745774	CF745774 UI-M-GV0-
	29	471.8	37.4	755	5	BU371858	BU371858 603568233
	30	469.8	37.3	794	5	BU205357	BU205357 603104569
	31	468.2	37.2	824	4	BI253163	BI253163 602973284
	32	467.2	37.1	785	6	CB990192	CB990192 AGENCOURT
	33	463.6	36.8	870	5	BQ216813	BQ216813 AGENCOURT
	34	463.4	36.8	484	5	BX281211	BX281211 BX281211
	35	463.4	36.8	699	5	BU707284	BU707284 UI-M-FR0-
	36	460	36.5	696	5	BU439278	BU439278 604147528
c	37	456.6	36.2	483	2	BF929411	BF929411 IL2-NT020
	38	454.2	36.0	631	5	BU112717	BU112717 603131421
	39	451.8	35.9	743	4	BI562440	BI562440 603256324
c	40	450.4	35.7	986	5	BU751496	BU751496 CH4#002_E
	41	449	35.6	933	2	BE886040	BE886040 601507192
	42	433.2	34.4	820	7	CK597557	CK597557 AGENCOURT
	43	427.4	33.9	672	5	BX261106	BX261106 BX261106
	44	420	33.3	604	5	BX261107	BX261107 BX261107
	45	413.2	32.8	570	5	BU293329	BU293329 604166812

ALIGNMENTS

RESULT 1
AK013347
LOCUS AK013347 1763 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810454O04 product:kinase interacting with leukemia-associated gene (stathmin), full insert sequence.
ACCESSION AK013347
VERSION AK013347.1 GI:12850651
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1763)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES Location/Qualifiers
source 1. .1763
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2810454O04"
/db_xref="taxon:10090"
/clone="2810454O04"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10, 11 days embryo"
CDS 159. .1418
/note="unnamed protein product; kinase interacting with leukemia-associated gene (stathmin) (MGD|MGI:1341908) putative"
/codon_start=1
/protein_id="BAB28802.1"
/db_xref="GI:12850652"
/translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALXQLQGHRNIVTLYGVFTIHFSP NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE CTSAVDLWSLGITLLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR MFDGKFVVATFYPLSAYKRGYLYQTLL"

ORIGIN

Query Match 88.4%; Score 1113.4; DB 3; Length 1763;
Best Local Similarity 92.7%; Pred. No. 5.4e-313;
Matches 1168; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

```
Qy       1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGCTTTTCTGGAGGCCTTCGGG 60
Db     159 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 218

Qy      61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
Db     219 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 278

Qy     121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
Db     279 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 338

Qy     181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
Db     339 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 398

Qy     241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
Db     399 GANCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 458

Qy     301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
Db     459 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 518

Qy     361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
Db     519 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 578

Qy     421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
Db     579 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 638

Qy     481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
Db     639 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 698

Qy     541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
Db     699 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 758

Qy     601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
Db     759 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 818

Qy     661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
Db     819 TCAGCTGTTGATCTGTGGAGCCTCGGAATCACTTTACTGGAAATGTTCTCAGGAATGAAG 878

Qy     721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
Db     879 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAT 938

Qy     781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
Db     939 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 998

Qy     841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
Db     999 ATCAAAAGCATGCTTCATGATGACCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGC 1058

Qy     901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
Db    1059 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACT 1118

Qy     961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
Db    1119 CCAGTGCTCAGGCTCCTCAATGTGCTGGATGATGATTACCTTGAAAATGAAGATGAATAT 1178

Qy    1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
Db    1179 GAAGATGTTGTAGAAGATGTGAAAGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1238

Qy    1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
```

```
                ||||||||||||||||||||||||||||||||||||||||| |||||||| || || ||||||
Db       1239 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1298

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
                ||||||||||||||| ||||| || |||||||| |||||||||| ||||||||||||||
Db       1299 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1358

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
                ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1359 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1418
```

RESULT 2
AK030152
LOCUS AK030152 2846 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4933401D07 product:kinase interacting with leukemia-associated gene (stathmin), full insert sequence.
ACCESSION AK030152
VERSION AK030152.1 GI:26326136
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 2846)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.

```
FEATURES             Location/Qualifiers
     source          1. .2846
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4933401D07"
                     /db_xref="taxon:10090"
                     /clone="4933401D07"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             238. .1230
                     /note="unnamed protein product; kinase interacting with
                     leukemia-associated gene (stathmin) (MGD|MGI:1341908)
                     putative"
                     /codon_start=1
                     /protein_id="BAC26812.1"
                     /db_xref="GI:26326137"
                     /translation="MTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMW
                     MIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY
                     IQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQE
                     WKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALCSPFFSIP
                     FAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSLLVPKE
                     NPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL"
     polyA_signal    2823. .2828
                     /note="putative"
     polyA_site      2846
                     /note="putative"
ORIGIN
```

Query Match 70.4%; Score 886.6; DB 3; Length 2846;
Best Local Similarity 93.1%; Pred. No. 8.8e-247;
Matches 928; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

```
Qy        264 CATCGTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTG 323
              | || |||||||| |||||||| ||||| || ||||| ||||| ||||||||||||||||||
Db        234 CTTCATGACTTTATATGGAGTCTTTACCATACACTTCTCTCCCAATGTGCCATCACGCTG 293

Qy        324 TCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCA 383
              |||| ||||||||||||||||||||||||||||||||||||| |||||||||||||| ||
Db        294 TCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTTTTATATTCCAGTCATCA 353

Qy        384 GGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGAGATGTTTTGGAGGCCCTTGCTTT 443
              ||| || |||||||||||||||||||| || |||||||||||| |||||||||| |||||
Db        354 GGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGAGATGTTCTGGAGGCCCTCGCTTT 413

Qy        444 TCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGC 503
              |||||| |||||||||| |||||||| |||||||||||||| ||||| ||||||||||
Db        414 TCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAACCACGAAACATCCTGTGGAGTGC 473

Qy        504 AGAGAATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGA 563
              |||||||| |||||||| || |||||||||||||| |||||||||||||||||||||||
Db        474 GGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTCAGCTTCAAAGAAGGCAATCAGGA 533
```

```
Qy        564 TGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTT 623
              ||||||||||||||||||||||||||||||| ||||| |||||||| |||| || |||||
Db        534 CGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCTGAAGCAGAGCTGCAGAACTGCTT 593

Qy        624 GGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCT 683
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        594 GGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCT 653

Qy        684 AGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCA 743
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        654 CGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAGCTGAAACATACAGTCAGATCTCA 713

Qy        744 GGAATGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGT 803
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        714 GGAATGGAAGGCAAACAGTTCTGCTATTATTGATCATATATTTGCCAGTAAAGCAGTGGT 773

Qy        804 GAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 863
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        774 GAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTTATCAAAAGCATGCTTCATGATGA 833

Qy        864 TCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTT 923
              ||  |||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db        834 CCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTT 893

Qy        924 TGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGT 983
              |||||||||||||||||||||||| |||||||| |||||||||| || || || |||||
Db        894 TGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACTCCAGTGCTCAGGCTCCTCAATGT 953

Qy        984 GCTGGATGATGATTATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAA 1043
              |||||||||||||| |||| |||||||| |||||||||||||||||||||||||||| ||
Db        954 GCTGGATGATGATTACCTTGAAAATGAAGATGAATATGAAGATGTTGTAGAAGATGTGAA 1013

Qy       1044 AGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGG 1103
              |||||||||||| |||||||||||||||||| ||||| |||||||||||||||||||||
Db       1014 AGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTGCTTGTTCCAAAGGAAAATCCTGG 1073

Qy       1104 CAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATT 1163
              |||||||||||||| |||||||| || || |||||||||||||||||||| ||||| ||
Db       1074 CAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGTGATTCCAAAGCTGCTCAGAAGTT 1133

Qy       1164 ACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGC 1223
              .|||||||| |||||||||| |||||||||||||||||||||||||||||||||||||||
Db       1134 GCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGC 1193

Qy       1224 CTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||| ||||||||||||||||||
Db       1194 CTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1230


RESULT 3
BM468107
LOCUS       BM468107               1079 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6432293 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5535653
            5', mRNA sequence.
ACCESSION   BM468107
VERSION     BM468107.1  GI:18517149
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12224  row: b  column: 06
            High quality sequence stop: 754.
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5535653"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

  Query Match             68.4%;  Score 861.8;  DB 4;  Length 1079;
  Best Local Similarity   98.8%;  Pred. No. 1.1e-239;
  Matches  888;  Conservative    0;  Mismatches    9;  Indels    2;  Gaps    2;

Qy        267 CGTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCT 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCT 60

Qy        327 GTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGG 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGG 120

Qy        387 TTGTTCCATGTGGATGATACAGCATTGCGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCT 446
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        121 TTGTTCCATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCT 180

Qy        447 TCATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGA 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGA 240

Qy        507 GAATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGT 566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGT 300

Qy        567 AAAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGC 626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGC 360

Qy        627 CCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGG 686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGG 420

Qy        687 AATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGA 746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGA 480

Qy        747 ATGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAA 806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAA 540

Qy        807 TGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCC 866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCC 600

Qy        867 AAGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGC 926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGC 660

Qy        927 CCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCT 986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCT 720

Qy        987 GGATGATGATTATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGA 1046
              |||||||||||||||| |||||||||||||||||||||||||||||||||    |||||
Db        721 GGATGATGATTATCTTGAGAATGAAGAGGAATATGAAGATGTTGTAGAAGAATGTAAAGA 780
```

```
Qy       1047 GGAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAG 1106
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        781 GGAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGNTCCAAAGGAAAATCCTGGCAG 840

Qy       1107 AGGACAAGTC-TTTGTTGAGTATGCAAATGCTG-GTGATTCCAAAGCTGCGCAGAAATT 1163
              |||||||||| |||| |||||||||||||||| ||||||| ||||||||||||||||||
Db        841 AGGACAAGTCTTTTGNTGAGTATGCAAATGCTGNGTGATTCAAAAGCTGCGCAGAAATT 899


RESULT 4
BM451184
LOCUS       BM451184                 917 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6392719 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5495159
            5', mRNA sequence.
ACCESSION   BM451184
VERSION     BM451184.1  GI:18500224
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12121  row: j  column: 24
            High quality sequence stop: 718.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5495159"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             57.4%;  Score 723.8;  DB 4;  Length 917;
  Best Local Similarity   99.7%;  Pred. No. 2e-199;
  Matches  725;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        534 TGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCG 593
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 TGGGCTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCG 80

Qy        594 GGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGA 653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 GGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGA 140

Qy        654 ATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 200

Qy        714 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 260

Qy        774 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 320

Qy        834 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 380

Qy        894 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 440

Qy        954 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGA 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        441 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGA 500

Qy       1014 GGAATATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGT 1073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 GGAATATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGT 560

Qy       1074 ATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAA 1133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 ATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAA 620

Qy       1134 TGCTGGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTT 1193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 TGCTGGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTT 680

Qy       1194 TGTTGTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTT 1253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 TGTTGTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTT 740

Qy       1254 GCTTTAA 1260
              |||||||
Db        741 GCTTTAA 747


RESULT 5
BY712344
LOCUS       BY712344                 1015 bp    mRNA    linear   EST 17-DEC-2002
DEFINITION  BY712344 RIKEN full-length enriched, 10, 11 days embryo whole body
            Mus musculus cDNA clone 2810454O04 5', mRNA sequence.
ACCESSION   BY712344
VERSION     BY712344.1  GI:27123621
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1015)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
```

```
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
  MEDLINE   22354683
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
            Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
            Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
            Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
            Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
            Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .1015
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="2810454O04"
                     /tissue_type="whole body"
                     /dev_stage="10, 11 days embryo"
                     /clone_lib="RIKEN full-length enriched, 10, 11 days embryo
                     whole body"
ORIGIN

  Query Match             49.4%;  Score 622.8;  DB 6;  Length 1015;
  Best Local Similarity   90.0%;  Pred. No. 6.3e-170;
  Matches  678;  Conservative    0;  Mismatches   73;  Indels    2;  Gaps    1;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              |||||||| |||||||||| |||||||| |||||||||| || ||||||||||||||||
Db        159 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 218

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              |||||||||||||| |||||||| ||||| |||||||| || |||||||||| |||||
Db        219 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 278

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              || |||||||| | ||| |||||||||||||||||||| ||||||| |||| || |||
Db        279 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 338

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||

Db        339 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 398

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              || |||||||||||||||| |||||||||||| |||||||| ||||| || |||||
Db        399 GANCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 458

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        459 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 518

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||| |||||||||||||| ||||| || |||||||||||||||||||| || ||||||
Db        519 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 578

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||| |||||||||| |||||||||| |||||||||| |||||||| |||||||||
Db        579 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 638

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||| |||||  |||||||||| |||||||| |||||||| || ||||||||||||||
Db        639 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 698

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||| ||||||||||||||||||||||||||||| |||||
Db        699 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 758

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              |||||||| |||| || |||||||||||||| ||||||||||||||||||||||||||||
Db        759 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 818

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||||| | | || ||||||| |||||||| ||||||||||||||||||
Db        819 TCAGCTGTTGATCTGTCGGGGCTCGGAATCACTTTACTGGAGATGTTCTCAGGAATGAAG 878

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAG 753
              |||||  |     |||||||||||||  ||| |
Db        879 CTGAA--AATACATCAGATCTCAGGATGGGAGG 909


RESULT 6
BG547675
LOCUS       BG547675                657 bp    mRNA    linear   EST 04-APR-2001
DEFINITION  602575520F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703645 5',
            mRNA sequence.
ACCESSION   BG547675
VERSION     BG547675.1  GI:13546340
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1542  row: g  column: 06
            High quality sequence stop: 657.
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4703645"
                     /lab_host="DH10B (T1 phage-resistant)"
```

/clone_lib="NIH_MGC_77"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 49.1%; Score 619.2; DB 4; Length 657;
Best Local Similarity 99.1%; Pred. No. 6.2e-169;
Matches 654; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

```
Qy        577 CAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGC 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGC 60

Qy        637 CTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTA 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTA 120

Qy        697 CTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCA 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCA 180

Qy        757 AACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATT 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATT 240

Qy        817 CCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGA 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGA 300

Qy        877 ATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATT 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATT 360

Qy        937 GAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGAT 996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGAT 420

Qy        997 TATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAA 1056
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TATCTTGAGAATGAAGAGGAATATGAAGATGTCGTAGAAGATGTAAAAGAGGAGTGTCAA 480

Qy       1057 AAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTC 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTC 540

Qy       1117 TTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGG 1176
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTTG-TGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGG 599

Qy       1177 ATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGA 1236
              |||||||||||||||| | ||||||||||||||||||||||||||| ||||||||||||||
Db        600 ATGTTTGATGGGAAGTTAG-TGTGGCTACATTCTACCCGCTGAGTG-CTACAAGAGGGGA 657
```

RESULT 7
BE895119
LOCUS BE895119 899 bp mRNA linear EST 20-OCT-2000
DEFINITION 601436041F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920939 5', mRNA sequence.
ACCESSION BE895119
VERSION BE895119.1 GI:10358193
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 899)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM9753 row: b column: 12
High quality sequence stop: 626.
FEATURES Location/Qualifiers
source 1. .899
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3920939"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 49.0%; Score 617.8; DB 2; Length 899;
Best Local Similarity 99.7%; Pred. No. 1.8e-168;
Matches 619; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        337 CTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATG 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATG 60

Qy        397 TGGATGATACAGCATTGCGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAG 456
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAG 120

Qy        457 GGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGT 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGT 180

Qy        517 TTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATT 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATT 240

Qy        577 CAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGC 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGC 300

Qy        637 CTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTA 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTA 360

Qy        697 CTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCA 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCA 420

Qy        757 AACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATT 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATT 480

Qy        817 CCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGA 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGA 540

Qy        877 ATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATT 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        541 ATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATT 600

Qy        937 GAAGATCTGGTCATGCTTCCC 957
              ||||||||||||||||||  ||
Db        601 GAAGATCTGGTCATGCTTTCC 621


RESULT 8
BI562420
LOCUS       BI562420                772 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603256316F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
            mRNA sequence.
ACCESSION   BI562420
VERSION     BI562420.1  GI:15449747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 772)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11755  row: p  column: 03
            High quality sequence stop: 727.
FEATURES             Location/Qualifiers
     source          1. .772
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5298674"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTTVN-3', size-selected for  average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation).  Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN

  Query Match             48.9%;  Score 616.4;  DB 4;  Length 772;
  Best Local Similarity   99.4%;  Pred. No. 4.3e-168;
  Matches  628;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        274 TTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTGTTGCTT 333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 TTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTGTTGCTT 167

Qy        334 GAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCC 393
              |||||||||||||||||||||||| |||||||||||||||||||| |||||||||
Db        168 GAACTCCTGGATGTCAGTGTTTCGGAANTGCTCTTATATTCCAGTCACCANGGTTGTTCC 227

Qy        394 ATGTGGATGATACAGCATTGCGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCAT 453
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        228 ATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCAT 287

Qy        454 GAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAA 513
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        288 GAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAA 347

Qy        514 TGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 573
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 TGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 407

Qy        574 ATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCT 633
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 ATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCT 467

Qy        634 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 693
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 527

Qy        694 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 587

Qy        754 GCAAACAGTTCTGCTATTATTGATCACA-TATTTGCCAGTAAAGCAGTGGTGAATGCCGC 812
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        588 GCAAACAGTTCTGCTATTATTGATCACATTATTTGCCAGTAAAGCAGTGGTGAATGCCGC 647

Qy        813 AATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAG 872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 AATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAG 707

Qy        873 AAGAATTCCTGCTGAAATGGCATTGTGCAGCC 904
              ||||||||||||||||||||||||||||||||
Db        708 AAGAATTCCTGCTGAAATGGCATTGTGCAGCC 739


RESULT 9
BU541306
LOCUS       BU541306                932 bp    mRNA    linear   EST 13-SEP-2002
DEFINITION  AGENCOURT_10327419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6572970
            5', mRNA sequence.
ACCESSION   BU541306
VERSION     BU541306.1  GI:22851747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 932)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2765  row: o  column: 18
            High quality sequence stop: 631.
FEATURES             Location/Qualifiers
     source          1. .932
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6572970"
                     /tissue_type="carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_40"
                     /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed  by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California,  Berkeley) using ZAP-cDNA synthesis kit
```

(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
ORIGIN

Query Match 48.3%; Score 608.8; DB 5; Length 932;
Best Local Similarity 93.7%; Pred. No. 7.5e-166;
Matches 689; Conservative 0; Mismatches 38; Indels 8; Gaps 5;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 208

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 268

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        269 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 328

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 388

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 448

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 508

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||| |||||||||||||||||||||||||||||| ||||||
Db        509 TTGCTCTTATATTCCAGTCACCAGCGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 568

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCT-CAA 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        569 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCCAA 628

Qy        480 ACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGAC- 538
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        629 ACCACGTAACATATTGTGGAGTGCAAAGAATGAATGTTTTAAACTCATTGACTTTGGACT 688

Qy        539 TTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTC 598
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        689 TTAGCTTCAAAGAAAGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTC 748

Qy        599 CAG-AAGCAGAATTGCAAAATTGCTTGGCCCAGGC-TGGCCTGCAGAGTGATACAGAATG 656
              ||| || |||||||||||||||||||||||||||| || ||||||||| |||||||| ||||
Db        749 CAGAAAACAGAATTGCAAAATTGCTTGGCCCANGCTTGGCCTGCAAAGTGATACAAAATG 808

Qy        657 TACCT----CAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAG 712
              |||||    | | || |  | ||  |  |  || |||||||||| | ||  |  ||
Db        809 TACCTCCAGCTGTTGATCTTGTGGAGCCCTAAGAAATCATTTTACCTGGAAAATGGTCTC 868

Qy        713 GAATGAAACTGAAAC 727
               |  |||    ||||
Db        869 CAGGGAATGGAAAAC 883
```

RESULT 10
BI259568
LOCUS BI259568 875 bp mRNA linear EST 17-JUL-2001
DEFINITION 602968330F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5108094 5',
mRNA sequence.
ACCESSION BI259568
VERSION BI259568.1 GI:14817032
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 875)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11261 row: k column: 07
High quality sequence stop: 722.
FEATURES Location/Qualifiers
source 1. .875
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5108094"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
ORIGIN

Query Match 47.2%; Score 595; DB 4; Length 875;
Best Local Similarity 91.6%; Pred. No. 7.8e-162;
Matches 809; Conservative 0; Mismatches 50; Indels 24; Gaps 16;

```
Qy        297 CTTTTCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTC 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db          1 CTTTTCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAG-GTTTC 59

Qy        357 GGAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGC 416
              |||||||||||||||||||||||||| |||||||||||||||||||||||||| || ||
Db         60 GGAATTGCTCTTATATTCCAGTCACCA-GGTTGTTCCATGTGGATGATACAGCA-TGTGC 117

Qy        417 CCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCT 476
              |||||||||||  |  | || ||||||||||||||||||||||||||||||||||||||
Db        118 CCGAGATGTTT--GGAGGCCCTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCT 175

Qy        477 CAAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGG 536
              |||||||||||||||| |||||||||||||||||||||||||||||||||||| ||||
Db        176 CAAACCACGTAACATA-TGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGAC-TTGG 233

Qy        537 ACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGC 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGC 293

Qy        597 TCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATG 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 TCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATG 353

Qy        657 TACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAAT 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        354 TACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAAT 413

Qy        717 GAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGA 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 GAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGA 473

Qy        777 TCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGA 836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 TCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGA 533

Qy        837 CCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATT 896
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        534 CCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCA-T 592
```

```
Qy        897 GTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCT-GGTCATGCTTC 955
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        593 GTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGGTCATGCTTC 652

Qy        956 CCACTCCAGTG--CTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAA-G 1012
              ||||||||| |    |||||||||||||||||||||||||||||||||| ||||||| |
Db        653 CCACTCCAGGGGCTAAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAACG 712

Qy       1013 AGGAATATGAAG-ATGTTGTAGAAG--ATGTAAAAGAGGAGTGTCAAAAATATGGACCAG 1069
              |||||||||| | |||||||||| |      ||| ||||||||||||  |  | |  |
Db        713 AGGAATATGACGAATGTTGTAGACGGATGTAAAAGGAGGAGTGTCAACCAATTTGGACCA 772

Qy       1070 TGGTATCTCTACTTGTT------CCAAAGGAAAATCCTGGCAGAGGACA-AGTCTTTGTT 1122
               |   | |   ||  ||        || | ||||||||||  ||||||| |||||||||
Db        773 GGTGGTATTCTCTACTTTGTTCCAAAAGGAAAAATCCTGGGCGAGGACACAGTCTTTGTA 832

Qy       1123 G-AGTATGCAAAT-GCTGGTGATTCCAAAGCTGCGCAGAAATT 1163
              | | |||||||| | ||| |||||||||  |||||  || ||
Db        833 GAAATATGCAAATGGTTGGGGATTCCAAAAGTGCGCGAAATTT 875
```

RESULT 11
CF593363
LOCUS CF593363 834 bp mRNA linear EST 26-SEP-2003
DEFINITION AGENCOURT_15622279 NIH_MGC_147 Homo sapiens cDNA clone IMAGE:30531268 5', mRNA sequence.
ACCESSION CF593363
VERSION CF593363.1 GI:36346826
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 834)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: NDAM621 row: m column: 05
High quality sequence stop: 660.
FEATURES Location/Qualifiers
source 1. .834
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30531268"
/tissue_type="Human Placenta"
/lab_host="DH10B TonA"
/clone_lib="NIH_MGC_147"
/note="Organ: placenta; Vector: pBluescriptR; Site_1: alI-XhoI; Site_2: BamH; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: This is a NIH_MGC library."
ORIGIN

Query Match 47.0%; Score 591.6; DB 7; Length 834;
Best Local Similarity 97.4%; Pred. No. 7.6e-161;
Matches 644; Conservative 0; Mismatches 10; Indels 7; Gaps 4;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 214

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 274

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        275 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 334

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 394

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 454

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 514

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        515 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 574

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        575 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 634

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        635 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTANACTCATTGACTTTGGACTT 694

Qy        541 A-GCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATC-GGGCTC 598
              | |||||||||||| ||||||||||||||||||||||||||||||||||||| ||||||
Db        695 AGGCTTCAAAGAAAGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGGCTC 754

Qy        599 CAGAAGCAGAATTGCAAAATTGCTT---GGCCCAGGCT--GGCCTGCAGAGTGATACAGA 653
              |||||||||||||||||||| |   | ||||||||  | ||||||||||||||||||||
Db        755 CAGAAGCAGAATTGCAAAAATTGCTTGGGCCCCAGGCTTGGCCCTGCAGAGTGATACAGA 814

Qy        654 A 654
              |
Db        815 A 815
```

RESULT 12
BG386905
LOCUS BG386905 906 bp mRNA linear EST 12-MAR-2001
DEFINITION 602454704F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4583065 5', mRNA sequence.
ACCESSION BG386905
VERSION BG386905.1 GI:13280454
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 906)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory

```
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: NIH Intramural Sequencing Center
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM1306  row: g  column: 02
           High quality sequence stop: 668.
FEATURES             Location/Qualifiers
     source          1. .906
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4583065"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_15"
                     /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)"
ORIGIN

  Query Match             46.3%;  Score 583.8;  DB 4;  Length 906;
  Best Local Similarity   96.1%;  Pred. No. 1.5e-158;
  Matches  664;  Conservative    0;  Mismatches   17;  Indels   10;  Gaps    6;

Qy        570 GTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCA 629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCA 61

Qy        630 GGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAAT 689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAAT 121

Qy        690 CATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATG 749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATG 181

Qy        750 GAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGC 809
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGC 241

Qy        810 CGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAG 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAG 301

Qy        870 CAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCC 929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCC 361

Qy        930 TCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGA 989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGA 421

Qy        990 TGATGATTATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGAGGA 1049
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        422 TGATGATTATCTTGAGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGAGGA 481

Qy       1050 GTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGG 1109
              |||||||||||||||||||||||||||||| |||||||||||||||||||||| |||||
Db        482 GTGTCAAAAATATGGACCAGTGGTATCTCTA-TTGTTCCAAAGGAAAATCCTGG-AGAGG 539

Qy       1110 ACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATTACTGAC 1169
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        540 ACAAGTCTTTGTTGAGTATGCAAATG-TGGTGATTCCAAAGCTGCGCAGAAATTACTGAC 598

Qy       1170 TGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGCCTACAA 1229
              ||||||    |||||||||||||   |   | ||||||||||| |||||||    |
Db        599 TGGAAG--ATGTTGATGGGAAGTTGTGT---GGTACATTCTACCCCGTGAGTGC--TACA 651

Qy       1230 GAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||| |||||||
Db        652 AGAGGGATATCTGTATCAAACCTGGCTTTAA 682


RESULT 13
BI333581
LOCUS       BI333581                 682 bp    mRNA    linear   EST 30-JUL-2001
DEFINITION  602997069F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5139184 5',
            mRNA sequence.
ACCESSION   BI333581
VERSION     BI333581.1  GI:15018238
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM11342  row: j  column: 17
           High quality sequence stop: 674.
FEATURES             Location/Qualifiers
     source          1. .682
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5139184"
                     /tissue_type="cervical carcinoma cell line"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_12"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.4 kb. Library prepared by Life
                     Technologies."
ORIGIN

  Query Match             46.3%;  Score 583.6;  DB 4;  Length 682;
  Best Local Similarity   97.1%;  Pred. No. 1.5e-158;
  Matches  647;  Conservative    0;  Mismatches   14;  Indels    5;  Gaps    5;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 78

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 138

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        139 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 198

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 258

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 318

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
```

```
Db        319 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 378

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420

Db        379 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 438

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Db        439 GATG-TTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 497

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Db        498 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 557

Qy        541 AGCTTCAAAGAAGGCAATCA-GGATGTAAAGTATATTCAGACAG-ACGGGTATCGGGCTC 598

Db        558 AGCTTCAAAGAAGGCAATCACGGATGTAAAGTATATTCAGACAGAACGGGTATCGGGCTC 617

Qy        599 CAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATA-CAGAATGT 657

Db        618 CAAGAAGCAGAATTGCAAAATGCTTGG-CCAGGCTGGCCTGCAGAGTGATACCAGAATGT 676

Qy        658 ACCTCA 663

Db        677 ACCTCA 682


RESULT 14
BQ225071
LOCUS       BQ225071                888 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7323233 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6016796
            5', mRNA sequence.
ACCESSION   BQ225071
VERSION     BQ225071.1  GI:20406471
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 888)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13215  row: a  column: 21
            High quality sequence stop: 581.
FEATURES             Location/Qualifiers
     source          1. .888
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6016796"
                     /tissue_type="epithelioid carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_70"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.1 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             45.9%;  Score 578;  DB 5;  Length 888;
  Best Local Similarity   97.8%;  Pred. No. 7.2e-157;
  Matches  618;  Conservative    0;  Mismatches   10;  Indels    4;  Gaps    3;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Db        149 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 208

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Db        209 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 268

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180

Db        269 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 328

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Db        329 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 388

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Db        389 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 448

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Db        449 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 508

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420

Db        509 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 568

Qy        421 GATGTTTTGGAGG-CCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAA 479

Db        569 GATGTTTTGGAGGCCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAA 628

Qy        480 ACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACT 539

Db        629 ACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACT 688

Qy        540 TAGCTTCAAAGAA-GGCAATCAGGATGTAAAGTATATTCAGACAGACGGGT--ATCGGGC 596

Db        689 TAGCTTCAAAGAAGGGAAATCAAGATGTAAAGTATATTCAGACCGACGGGTTATCGGGGC 748

Qy        597 TCCAGAAGCAGAATTGCAAAATTGCTTGGCCC 628

Db        749 TCCAGAAGCAAAATTGCAAAATTGCTTGGGCC 780


RESULT 15
CO737403
LOCUS       CO737403                602 bp    mRNA    linear   EST 29-JUL-2004
DEFINITION  SlLT02c17i16f1 squirrel testis library 1 Spermophilus lateralis
            cDNA clone 17i16 5', mRNA sequence.
ACCESSION   CO737403
VERSION     CO737403.1  GI:50824673
KEYWORDS    EST.
SOURCE      Spermophilus lateralis (golden-mantled ground squirrel)
  ORGANISM  Spermophilus lateralis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
            Spermophilus.
REFERENCE   1  (bases 1 to 602)
  AUTHORS   Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
            Rogers,J. and Cossins,A.R.
  TITLE     Microarray analysis of transcriptional changes during hibernation
            in the golden mantled ground squirrel, Spermophilus lateralis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Andrew R. Cossins
            Laboratory for Environmental Gene Regulation
            University of Liverpool
            School of Biological Sciences, The Biosciences Building, Crown
            Street, Liverpool, United Kingdom, L69 7ZB
            Tel: +44(0)151-795-4510
            Fax: +44(0)151-795-4431
            Email: cossins@liv.ac.uk
```

```
            Vector has been trimmed from this EST.
            Plate: 17, row: i  column: 16
            Seq primer: pflc T7 (5'-AATACGACTCACTATAGGG-3')
            High quality sequence stop: 602.
FEATURES              Location/Qualifiers
     source           1. .602
                      /organism="Spermophilus lateralis"
                      /mol_type="mRNA"
                      /db_xref="taxon:76772"
                      /clone="17i16"
                      /sex="Male & female"
                      /tissue_type="Testis"
                      /dev_stage="Adult"
                      /lab_host="E.coli Electromax DH10B"
                      /clone_lib="squirrel testis library 1"
                      /note="Vector: pFLC; Site_1: SalI GTCGAG; Site_2: BamHI
                      GGATCC; Normalized and subtracted cDNA library prepared
                      from testis of hibernating and summer animals"
ORIGIN

  Query Match             44.4%;  Score 560;  DB 7;  Length 602;
  Best Local Similarity   95.8%;  Pred. No. 1.1e-151;
  Matches  575;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy        624 GGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCT 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCT 62

Qy        684 AGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCA 743
              |||||| ||||||||||||||||||||||||||||||||||||||| ||||||||
Db         63 AGGAATTATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTTAGATCTCA 122

Qy        744 GGAATGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGT 803
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGAATGGAAGGCAAATAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGT 182

Qy        804 GAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 242

Qy        864 TCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTT 923
              |||||||||||| |||||||||||||||||||||||| |||||||| ||||||||||||
Db        243 TCCAAGCAGAAGGATTCCTGCTGAAATGGCATTGTGCAGTCCATTCTTCAGCATTCCTTT 302

Qy        924 TGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGT 983
              |||||||||||||||||| |||| |||||||| |||||||| |||||| |||| |||||
Db        303 TGCCCCTCATATTGAAGATTTGGTGATGCTTCCTACTCCAGTCCTAAGATTGCTCAATGT 362

Qy        984 GCTGGATGATGATTATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAA 1043
              |||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||
Db        363 CCTGGATGATGATTATCTTGAGAATGAAGAAGAATATGAAGATGTTGTAGAAGATGTAAA 422

Qy       1044 AGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGG 1103
              |||||| |||||||||||||||||| || ||||| ||||||||||||||||||||||||
Db        423 AGAGGAATGTCAAAAATATGGACCAGTAGTGTCTCTGCTTGTTCCAAAGGAAAATCCTGG 482

Qy       1104 CAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATT 1163
              |||||| |||||||||||||||||||||| |||||||||||||||||| ||||||||
Db        483 CAGAGGGCAAGTCTTTGTTGAGTATGCAAATCCTGGTGATTCCAAAGCTGCTCAGAAATT 542

Qy       1164 ACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGC 1223
              ||| |||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db        543 ACTAACTGGAAGGATGTTTGATGGGAAATTCGTTGTGGCTACATTCTACCCGCTGAGTGC 602


Search completed: November 29, 2004, 17:46:26
Job time : 5313 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:30 ; Search time 5510 Seconds
(without alignments)
10813.978 Million cell updates/sec

Title: US-10-798-532-3
Perfect score: 1260
Sequence: 1 atggcgggatccggctgcgc.....tgtatcaaaccttgctttaa 1260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1255.2	99.6	1260	6	CQ715025	CQ715025 Sequence
2	1255.2	99.6	1260	6	AX262521	AX262521 Sequence
3	1255.2	99.6	2598	6	AX262519	AX262519 Sequence
4	1255.2	99.6	2909	9	HSA536197	AJ536197 Homo sapi
5	1255.2	99.6	4065	6	CQ412057	CQ412057 Sequence
6	1255.2	99.6	4065	6	AX188296	AX188296 Sequence
7	1253.6	99.5	1260	6	AX166543	AX166543 Sequence
8	1253.6	99.5	1824	6	AX680136	AX680136 Sequence
9	1114.4	88.4	2041	10	BC058732	BC058732 Mus muscu
c 10	1111.2	88.2	110000	2	AC129141_1	Continuation (2 of
11	1106.4	87.8	1633	10	RNKISRNA	X98374 R.norvegicu
12	1106.4	87.8	3244	10	RNU70372	U70372 Rattus norv
13	1098.4	87.2	1254	10	MMPKKIS	Y10725 Mus musculu
14	1065.6	84.6	1367	9	BC026046	BC026046 Homo sapi
15	989.8	78.6	1135	9	AK058195	AK058195 Homo sapi
16	941.8	74.7	2041	9	BC014917	BC014917 Homo sapi
17	768	61.0	1157	5	CR386837	CR386837 Gallus ga
18	699.8	55.5	2940	9	HSM805327	AL834136 Homo sapi
c 19	567.6	45.0	259562	2	AC094499	AC094499 Rattus no
20	535.4	42.5	601	10	MMKISRNA	X82320 M.musculus
21	458.6	36.4	529	6	AX186343	AX186343 Sequence
22	445.8	35.4	722	6	BD221413	BD221413 Human gen
23	426.6	33.9	470	6	AX187546	AX187546 Sequence
24	387.4	30.7	462	6	AX184680	AX184680 Sequence
25	295	23.4	105902	9	AL359699	AL359699 Human DNA
26	295	23.4	167995	2	AC040901	AC040901 Homo sapi
c 27	293.4	23.3	211812	2	AC027205	AC027205 Homo sapi
c 28	254.4	20.2	200095	10	AC121606	AC121606 Mus muscu
29	254.4	20.2	220806	10	AC123650	AC123650 Mus muscu
30	252.8	20.1	100207	10	AY180177	AY180177 Mus muscu
31	249.6	19.8	211635	2	AC112763	AC112763 Rattus no
32	203.8	16.2	231	6	BD220137	BD220137 Human gen
33	184.8	14.7	185810	2	AC136052	AC136052 Rattus no
34	177.4	14.1	419	6	CQ393016	CQ393016 Sequence
35	177.4	14.1	419	6	CQ399411	CQ399411 Sequence
36	164.6	13.1	60640	10	AY180176	AY180176 Mus muscu
37	153.4	12.2	205752	2	AC119431	AC119431 Mus muscu
c 38	87.4	6.9	185810	2	AC136052	AC136052 Rattus no
39	85	6.7	198	6	CQ405796	CQ405796 Sequence
40	69.6	5.5	2108	8	NPL18351	Y18351 Nicotiana p
41	64.6	5.1	2047	8	BT013617	BT013617 Lycopersi
42	63.4	5.0	65	6	CQ555763	CQ555763 Sequence
43	60.6	4.8	1996	5	BC044032	BC044032 Xenopus l
44	59.8	4.7	2204	8	NPL18350	Y18350 Nicotiana p
45	55.8	4.4	1565	5	BC065869	BC065869 Danio rer

ALIGNMENTS

RESULT 1
CQ715025
LOCUS CQ715025 1260 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 959 from Patent WO02068579.
ACCESSION CQ715025
VERSION CQ715025.1 GI:42275882
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 959 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1. .1260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 99.6%; Score 1255.2; DB 6; Length 1260;
Best Local Similarity 99.8%; Pred. No. 1.2e-310;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
```

```
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

RESULT 2

```
AX262521
LOCUS       AX262521                1260 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 12 from Patent WO0173050.
ACCESSION   AX262521
VERSION     AX262521.1  GI:16511408
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Meyers,R.
  TITLE     3714, 16742, 23546, and 13887 novel protein kinase molecules and
            uses therefor
  JOURNAL   Patent: WO 0173050-A 12 04-OCT-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1260
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           99.6%;  Score 1255.2;  DB 6;  Length 1260;
  Best Local Similarity 99.8%;  Pred. No. 1.2e-310;
  Matches 1257;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260


RESULT 3
AX262519
LOCUS       AX262519                2598 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 10 from Patent WO0173050.
ACCESSION   AX262519
VERSION     AX262519.1  GI:16511406
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Meyers,R.
  TITLE     3714, 16742, 23546, and 13887 novel protein kinase molecules and
            uses therefor
  JOURNAL   Patent: WO 0173050-A 10 04-OCT-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2598
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             268. .1527
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD10238.1"
                     /db_xref="GI:16511407"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPPFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

  Query Match           99.6%;  Score 1255.2;  DB 6;  Length 2598;
  Best Local Similarity 99.8%;  Pred. No. 1.4e-310;
  Matches 1257;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
```

```
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1527


RESULT 4
HSA536197
LOCUS       HSA536197               2909 bp    mRNA    linear   PRI 11-JAN-2003
DEFINITION  Homo sapiens mRNA for KIS protein.
ACCESSION   AJ536197
VERSION     AJ536197.1  GI:27657360
KEYWORDS    KIS protein.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Bieche,I., Manceau,V., Curmi,P.A., Laurendon,I., Lachkar,S.,
            Leroy,K., Vidaud,D., Sobel,A. and Maucuer,A.
  TITLE     Quantitative RT-PCR reveals a ubiquitous but preferentially neural
            expression of the KIS gene in rat and human
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2909)
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-2003) Maucuer A., U440, Inserm, IFM, 17 rue du
            fer a moulin, 75005 Paris, FRANCE
FEATURES             Location/Qualifiers
     source          1. .2909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:IMAGp998M213680"
                     /db_xref="taxon:9606"
                     /clone="I.M.A.G.E. Consortium CloneID 1455548"
                     /tissue_type="2 pooled tumors (clear cell type)"
     gene            1. .2909
                     /gene="KIS"
     CDS             146. .1405
                     /gene="KIS"
                     /codon_start=1
                     /product="KIS protein"
                     /protein_id="CAD60192.1"
                     /db_xref="GI:27657361"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

  Query Match           99.6%;  Score 1255.2;  DB 9;  Length 2909;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-310;
  Matches 1257;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        146 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 205

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 265

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        266 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 325

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 385

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 445

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 505

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        506 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 565

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 625

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 685

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 745

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 805

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 865

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 925

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 985

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1045

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1105

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1106 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1165

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1225

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1285
```

```
Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1345

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1346 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1405


RESULT 5
CQ412057
LOCUS       CQ412057                4065 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 19128 from Patent WO0170979.
ACCESSION   CQ412057
VERSION     CQ412057.1  GI:41319838
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Lillie,J.
  TITLE     Genes, compositions, kits, and method for identification,
            assessment, prevention, and therapy of ovarian cancer
  JOURNAL   Patent: WO 0170979-A 19128 27-SEP-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .4065
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             99.6%;  Score 1255.2;  DB 6;  Length 4065;
  Best Local Similarity   99.8%;  Pred. No. 1.5e-310;
  Matches 1257;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1527


RESULT 6
AX188296
LOCUS       AX188296                4065 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 3991 from Patent WO0142467.
ACCESSION   AX188296
VERSION     AX188296.1  GI:15139769
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
  TITLE     Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
  JOURNAL   Patent: WO 0142467-A 3991 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .4065
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match            99.6%;  Score 1255.2;  DB 6;  Length 4065;
  Best Local Similarity  99.8%;  Pred. No. 1.5e-310;
  Matches 1257;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1527


RESULT 7
AX166543
LOCUS       AX166543               1260 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 34 from Patent WO0138503.
ACCESSION   AX166543
VERSION     AX166543.1  GI:14546888
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
            Flanagan,P. and Clary,D.S.
  TITLE     Novel human protein kinases and protein kinase-like enzymes
  JOURNAL   Patent: WO 0138503-A 34 31-MAY-2001;
            Sugen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1260
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            99.5%;  Score 1253.6;  DB 6;  Length 1260;
  Best Local Similarity  99.7%;  Pred. No. 3e-310;
  Matches 1256;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420
```

```
Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

```
RESULT 8
AX680136
LOCUS       AX680136               1824 bp    DNA     linear   PAT 29-MAR-2003
DEFINITION  Sequence 21 from Patent WO0181555.
ACCESSION   AX680136
VERSION     AX680136.1  GI:29369916
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Gururajan,R., Patterson,C., Recipon,S.A., Tribouley,C.M.,
            Hafalia,A., Khan,F.Y., Yue,H., Au-Young,J., Bandman,O.,
            Baughn,M.R., Borowsky,M.L., Burford,N., Burrill,J.D., Elliot,V.S.,
            Gandhi,A.R., Kearney,L., Lal,P., Lu,D.A., Lu,Y., Tang,Y.T.,
            Zingler,K.A., Griffin,J.A., Hillman,J.J., Marcus,G.A., Nguyen,D.B.,
            Policky,J., Ramkumar,J., Thornton,M., Walia,N.K. and Walsh,R.T.
  TITLE     Human kinases
  JOURNAL   Patent: WO 0181555-A 21 01-NOV-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1824
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 3107952CB1"
ORIGIN

  Query Match             99.5%;  Score 1253.6;  DB 6;  Length 1824;
  Best Local Similarity   99.7%;  Pred. No. 3.3e-310;
  Matches 1256;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 228

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 288

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        289 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 348

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 408

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 468

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 528

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        529 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 588

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 648

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 708

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 768

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 828

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 888

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 948

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1008
```

```
Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1068

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1128

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1129 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1188

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1189 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1248

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1249 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1308

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1309 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1368

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1369 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1428
```

RESULT 9
BC058732
LOCUS BC058732 2041 bp mRNA linear ROD 30-JUN-2004
DEFINITION Mus musculus kinase interacting with leukemia-associated gene (stathmin), mRNA (cDNA clone MGC:64770 IMAGE:6414877), complete cds.
ACCESSION BC058732
VERSION BC058732.1 GI:37194892
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2041)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 2041)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
Web site: http://genome.uiowa.edu
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A., Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K., Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J., Casavant,T., Soares,M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754447.
FEATURES Location/Qualifiers
source 1. .2041
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:64770 IMAGE:6414877"
/tissue_type="Brain, pool of enriched mouse brain tissues from 13.5,14.5,16.5 and 17.5 dpc"
/clone_lib="NIH_BMAP_FR0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
gene 1. .2041
/gene="Kist"
/note="synonym: KIS"
/db_xref="LocusID:16589"
/db_xref="MGI:1341908"
CDS 154. .1413
/gene="Kist"
/codon_start=1
/product="kinase interacting with leukemia-associated gene (stathmin)"
/protein_id="AAH58732.1"
/db_xref="GI:37194893"
/db_xref="LocusID:16589"
/db_xref="MGI:1341908"
/translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

```
Query Match            88.4%;  Score 1114.4;  DB 10;  Length 2041;
Best Local Similarity  92.8%;  Pred. No. 1.7e-274;
Matches 1169;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              |||||||| |||||||||| |||||||| |||||||||| || ||||||||||||||||
Db        154 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 213

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||| |||||||| ||||| |||||||| || |||||||||| |||||
Db        214 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 273

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              || |||||||| | ||| |||||||||||||||||||| ||||||| |||| || |||
Db        274 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 333

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||
```

```
Db        334 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 393

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              || |||||||||||||||| |||||||||||| |||||||| ||||| || |||||
Db        394 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 453

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        454 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 513

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||| |||||||||||||| ||||| || |||||||||||||||||||| || ||||||
Db        514 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 573

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||| |||||||||| |||||||||| |||||||||| |||||||| |||||||||
Db        574 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 633

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||| |||||  |||||||||| |||||||| |||||||  || ||||||||||||||
Db        634 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 693

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||| ||||||||||||||||||||||||||||| |||||
Db        694 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 753

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              |||||||| |||| || |||||||||||||| ||||||||||||||||||||||||||||
Db        754 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 813

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        814 TCAGCTGTTGATCTGTGGAGCCTCGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAG 873

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        874 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAT 933

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        934 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 993

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              |||||||||||||||||||||| ||  |||||||||| ||||||||  ||||||||||||
Db        994 ATCAAAAGCATGCTTCATGATGACCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGC 1053

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||| |||
Db       1054 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACT 1113

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||| || || || |||||||||||||||||||| ||||  |||||||| ||||||
Db       1114 CCAGTGCTCAGGCTCCTCAATGTGCTGGATGATGATTACCTTGAAAATGAAGATGAATAT 1173

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||| |||||
Db       1174 GAAGATGTTGTAGAAGATGTGAAAGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1233

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||||||||||||||||||||||||| |||||||| || || ||||||
Db       1234 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1293

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              |||||||||||||| ||||| || |||||||| |||||||||| ||||||||||||||||
Db       1294 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1353

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1354 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1413
```

RESULT 10
AC129141_1/c

```
WPCOMMENT
  Sequence split into 4 fragments  LOCUS AC129141 Accession AC129141
     Fragment Name      Begin     End
     AC129141_0             1    110000
     AC129141_1        100001    210000
     AC129141_2        200001    310000
     AC129141_3        300001    363206
Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH23

  Query Match           88.2%;  Score 1111.2;  DB 2;  Length 110000;
  Best Local Similarity   92.6%;  Pred. No. 2.5e-273;
  Matches 1167;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              |||||  || ||| |||||||| |||||||| |||||||||||  || ||||||||||||||
Db      70316 ATGGCCGGGTCCTGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 70257

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              |||||||||||||||||||||| || || |||||||||||  |||||||||  |||||
Db      70256 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 70197

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||  || |||||||||||||||||||||||||  |||||| |||| || |||
Db      70196 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 70137

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||| ||||| || || |||||||||||||| |||||||||| ||||||
Db      70136 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 70077

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              || |||||||||||||||| |||||||||||||| || ||||| ||||| || |||||
Db      70076 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 70017

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db      70016 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 69957

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||  |||||||||||||||||||| |||||||||||||||||||| || ||| ||
Db      69956 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATACAGCACTGTGCCAGA 69897

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||| |||||||||||||||||||| |||||||||||||||||||| |||||||||
Db      69896 GATGTTCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAA 69837

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||| |||||  |||||||||| ||||| || |||||  || ||||||||||||||
Db      69836 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 69777

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||| || |||||||||||||||||||||||||| |||||
Db      69776 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 69717

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              |||||||| |||| |||||||||||||||| |||||||||||||||||||| ||||||
Db      69716 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 69657

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||||| |||||||| |||||||||||||||||||||||||||||||||||
Db      69656 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 69597

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      69596 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 69537

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db      69536 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 69477

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              |||||||||||||||||||||| ||  |||||||||| ||||||||  ||||||||||||
Db      69476 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 69417
```

```
Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
Db      69416 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 69357

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
Db      69356 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 69297

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
Db      69296 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 69237

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
Db      69236 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 69177

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
Db      69176 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGATGGGAAGTTTGTTGTG 69117

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
Db      69116 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 69057


RESULT 11
RNKISRNA
LOCUS       RNKISRNA                1633 bp    mRNA    linear   ROD 08-JUL-2002
DEFINITION  R.norvegicus mRNA for KIS protein.
ACCESSION   X98374
VERSION     X98374.1  GI:1403531
KEYWORDS    kis gene.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Maucuer,A., Ozon,S., Manceau,V., Gavet,O., Lawler,S., Curmi,P. and
            Sobel,A.
  TITLE     KIS is a protein kinase with an RNA recognition motif
  JOURNAL   J. Biol. Chem. 272 (37), 23151-23156 (1997)
  MEDLINE   97435279
   PUBMED   9287318
REFERENCE   2  (bases 1 to 1633)
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-1996) A. Maucuer, INSERM U440, 17 rue du Fer
            Moulin, 75005 Paris, FRANCE
FEATURES             Location/Qualifiers
     source          1..1633
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /tissue_type="brain"
                     /clone_lib="7"
                     /dev_stage="neonate"
     gene            1..1633
                     /gene="KIS"
     CDS             39..1298
                     /gene="KIS"
                     /codon_start=1
                     /protein_id="CAA67021.1"
                     /db_xref="GI:1403532"
                     /db_xref="GOA:Q63285"
                     /db_xref="Swiss-Prot:Q63285"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

  Query Match             87.8%;  Score 1106.4;  DB 10;  Length 1633;
  Best Local Similarity   92.4%;  Pred. No. 1.8e-272;
  Matches 1164;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
Db         39 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 98

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
Db         99 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 158

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
Db        159 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 218

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
Db        219 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 278

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
Db        279 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 338

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
Db        339 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 398

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
Db        399 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 458

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
Db        459 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 518

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
Db        519 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 578

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
Db        579 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 638

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
Db        639 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 698

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
Db        699 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 758

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
Db        759 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 818

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
Db        819 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 878

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
Db        879 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 938

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
Db        939 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 998

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
```

```
Db        999 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1058

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              |||||||||||||||||||||||| |||||||| |||||||||||||||||| |||||
Db       1059 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1118

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db       1119 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1178

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              |||||||||||||| |||||||| |||||||| |||||||||| ||||||||||||||||
Db       1179 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1238

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1239 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1298


RESULT 12
RNU70372
LOCUS       RNU70372               3244 bp    mRNA    linear   ROD 08-SEP-1999
DEFINITION  Rattus norvegicus PAM COOH-terminal interactor protein 2 mRNA,
            complete cds.
ACCESSION   U70372
VERSION     U70372.2  GI:5821767
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 3244)
  AUTHORS   Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
            and Eipper,B.A.
  TITLE     Novel proteins that interact with the COOH-terminal cytosolic
            routing determinants of an integral membrane peptide-processing
            enzyme
  JOURNAL   J. Biol. Chem. 271 (45), 28636-28640 (1996)
  MEDLINE   97067094
   PUBMED   8910496
REFERENCE   2  (bases 1 to 3244)
  AUTHORS   Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
            Eipper,B.A. and Mains,R.E.
  TITLE     The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
            Determinants of the Peptide Processing Enzyme Peptidylglycine
            alpha-Amidating Monooxygenase
  JOURNAL   J. Biol. Chem. 274 (1999) In press
REFERENCE   3  (bases 1 to 3244)
  AUTHORS   Alam,R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
            and Eipper,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
            School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
            21205, USA
REFERENCE   4  (bases 1 to 3244)
  AUTHORS   Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
            Eipper,B.A. and Mains,R.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
            School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
            21205, USA
  REMARK    Sequence update by submitter
COMMENT     On Sep 3, 1999 this sequence version replaced gi:1698778.
FEATURES             Location/Qualifiers
     source          1. .3244
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
     CDS             162. .1421
                     /function="serine protein kinase"
                     /note="P-CIP2"
                     /codon_start=1
                     /product="PAM COOH-terminal interactor protein 2"
                     /protein_id="AAC53031.2"
                     /db_xref="GI:5821768"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

  Query Match           87.8%;  Score 1106.4;  DB 10;  Length 3244;
  Best Local Similarity   92.4%;  Pred. No. 2.1e-272;
  Matches 1164;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||| || |||||||||||| |||||||| |||||||||||| || ||||||||||||||
Db        162 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 221

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              |||||||||||||||||||||||||| || || |||||||||||| |||||||||| |||||
Db        222 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 281

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||   || |||||||||||||||||||| ||||||| |||| || |||
Db        282 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 341

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||| ||||| || || |||||||||||||| |||||||||| ||||||
Db        342 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 401

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              || |||||||||||||||||||||||||||||||||| || ||||| ||||| || |||||
Db        402 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 461

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        462 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 521

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||  |||||||||||||||||||| |||||||||||||| ||||| || ||| ||
Db        522 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 581

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||  |||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        582 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 641

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              |||||  |||||  |||||||||| |||||  || ||  |||||  || ||||||||||||||
Db        642 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 701

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||||| || |||||||||||||||||||||||||||| ||||||
Db        702 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 761

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              |||||||||  ||||  |||||||||||||||||||  |||||||||||||||||||| ||||||
Db        762 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 821

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||| |||||||| |||||||||||||||||||||||||||||||||||||
Db        822 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 881

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        882 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 941

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        942 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 1001
```

```
Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||| || |||||||||| || |||||||| ||||| ||||||
Db       1002 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 1061

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||
Db       1062 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 1121

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||| ||||| || |||||||||||||||| |||||||  |||||||| ||||||
Db       1122 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1181

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              |||||||||||||||||||||||| |||||||| |||||||||||||||| |||||
Db       1182 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1241

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db       1242 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1301

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              |||||||||||||| |||||||| |||||||| |||||||||| ||||||||||||||||
Db       1302 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1361

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1362 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1421


RESULT 13
MMPKKIS
LOCUS       MMPKKIS                 1254 bp    mRNA    linear   ROD 08-JUL-2002
DEFINITION  Mus musculus partial mRNA for KIS protein kinase (Kis gene).
ACCESSION   Y10725
VERSION     Y10725.2  GI:21726712
KEYWORDS    Kis gene; protein kinase.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1254)
  AUTHORS   Maucuer,A., Ozon,S., Manceau,V., Gavet,O., Lawler,S., Curmi,P. and
            Sobel,A.
  TITLE     KIS is a protein kinase with an RNA recognition motif
  JOURNAL   J. Biol. Chem. 272 (37), 23151-23156 (1997)
  MEDLINE   97435279
   PUBMED   9287318
REFERENCE   2
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-1997) Maucuer A., INSERM U440, U440, 17 Rue du
            Fer Moulin, Paris, 75005, FRANCE
  REMARK    revised by [4]
REFERENCE   3  (bases 1 to 1254)
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) A. Maucuer, INSERM U440, U440, 17 Rue du
            Fer Moulin, Paris, 75005, FRANCE
COMMENT     On Jul 10, 2002 this sequence version replaced gi:1806129.
            Overlaps with related sequences X82320, X10725.
FEATURES             Location/Qualifiers
     source          1. .1254
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="K22"
                     /tissue_type="brain"
                     /clone_lib="lambda gt10"
                     /dev_stage="neonate"
     gene            1. .1254
                     /gene="Kis"
     CDS             11. .>1254
                     /gene="Kis"
                     /codon_start=1
                     /product="KIS protein kinase"
                     /protein_id="CAA71714.2"
                     /db_xref="GI:21726713"
                     /db_xref="GOA:P97343"
                     /db_xref="Swiss-Prot:P97343"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYL"
ORIGIN

  Query Match           87.2%;  Score 1098.4;  DB 10;  Length 1254;
  Best Local Similarity   92.7%;  Pred. No. 1.9e-270;
  Matches 1153;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              |||||||| |||||||||| |||||||| |||||||||| || ||||||||||||||||
Db         11 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 70

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||| |||||||| ||||| |||||||| || |||||||||| |||||
Db         71 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 130

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              || |||||||| | ||| |||||||||||||||||||| ||||||| |||| || |||
Db        131 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 190

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||||||| ||||| |||||||||||||||||||||||||||||||||||
Db        191 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 250

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              || |||||||||||||||| |||||||||||| |||||||| ||||| || |||||
Db        251 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 310

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        311 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 370

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||| |||||||||||||| ||||| || |||||||||||||||||| || ||||||
Db        371 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 430

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||| |||||||||| |||||||||| |||||||||| |||||||| |||||||||
Db        431 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 490

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||| |||||  |||||||||| |||||||| |||||||| || ||||||||||||||
Db        491 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 550

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||| |||||
Db        551 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 610

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||  |||| || |||||||||||||| ||||||||||||||||||||||||||||
Db        611 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 670

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        671 TCAGCTGTTGATCTGTGGAGCCTCGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAG 730

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        731 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAT 790

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
```

```
Db        791 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 850

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||  ||  ||||||||||| ||||||| ||||||||||||
Db        851 ATCAAAAGCATGCTTCATGATGACCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGC 910

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||
Db        911 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACT 970

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||| || || || |||||||||||||||||||| ||||  |||||||| ||||||
Db        971 CCAGTGCTCAGGCTCCTCAATGTGCTGGATGATGATTACCTTGAAAATGAAGATGAATAT 1030

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              |||||||||||||||||||| |||||||||||||| |||||||||||||||||| |||||
Db       1031 GAAGATGTTGTAGAAGATGTGAAAGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1090

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||||||||||||||||||||||||| |||||||| || || ||||||
Db       1091 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1150

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              |||||||||||||| ||||| || |||||||| |||||||||||| ||||||||||||||||
Db       1151 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1210

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTA 1244
              |||||||||||||||||||||||||||||||||||||||||| ||
Db       1211 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTA 1254
```

RESULT 14
BC026046
LOCUS BC026046 1367 bp mRNA linear PRI 30-JUN-2004
DEFINITION Homo sapiens kinase interacting with leukemia-associated gene (stathmin), mRNA (cDNA clone MGC:39391 IMAGE:4649261), complete cds.
ACCESSION BC026046
VERSION BC026046.1 GI:19684094
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1367)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 1367)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu, Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 42 Row: f Column: 6
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21389438.

FEATURES Location/Qualifiers
source 1. .1367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:39391 IMAGE:4649261"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
gene 1. .1367
/gene="KIS"
/db_xref="LocusID:127933"
CDS 159. .1193
/gene="KIS"
/codon_start=1
/product="kinase interacting stathmin"
/protein_id="AAH26046.1"
/db_xref="GI:19684095"
/db_xref="LocusID:127933"
/translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSATIDHIFASKAVVNAAIPAYH
LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
ENEEEYEGLC"

ORIGIN

Query Match 84.6%; Score 1065.6; DB 9; Length 1367;
Best Local Similarity 92.6%; Pred. No. 5.1e-262;
Matches 1167; Conservative 0; Mismatches 4; Indels 89; Gaps 1;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 218

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 278

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        279 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 338

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 398
```

```
Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 458

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 518

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 578

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 638

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 698

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 758

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 818

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 878

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTACTATTGATCAC 938

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 998

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1058

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1118

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1119 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1178

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||
Db       1179 GAAG------------------------------------------------------- 1182

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
                                                ||||||||||||||||||||||||||||
Db       1183 -------------------------------GTCTTTGTTGAGTATGCAAATGCTGGT 1209

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1269

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1270 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1329
```

RESULT 15
AK058195
LOCUS AK058195 1135 bp mRNA linear PRI 12-SEP-2003

```
DEFINITION  Homo sapiens cDNA FLJ25466 fis, clone TST09304, highly similar to
            Kinase interacting with leukemia-associated gene (stathmin).
ACCESSION   AK058195
VERSION     AK058195.1  GI:16554273
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T.,
            Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R.,
            Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiyama,H.,
            Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E.,
            Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S.,
            Terashima,Y., Watanabe,M., Suzuki,Y., Hata,H., Nakagawa,K.,
            Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A., Kawakami,B.,
            Nagai,K., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1135)
  AUTHORS   Sugano,S. and Suzuki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure, Human Genome
            Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction and 5'-end one pass sequencing: Institute of Medical
            Science, University of Tokyo, Laboratory of Genome Structure, Human
            Genome Center; 3'-end one pass sequencing: RAB; clone selection for
            full insert sequencing: RAB and Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .1135
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="TST09304"
                     /tissue_type="testis"
                     /clone_lib="TST"
                     /note="cloning vector: pME18SFL3"
ORIGIN
```

Query Match 78.6%; Score 989.8; DB 9; Length 1135;
Best Local Similarity 99.8%; Pred. No. 1.3e-242;
Matches 991; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        268 GTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTG 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTG 182

Qy        328 TTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGT 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 TTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGT 242

Qy        388 TGTTCCATGTGGATGATACAGCATTGCGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTT 447
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        243 TGTTCCATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTT 302

Qy        448 CATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAG 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAG 362

Qy        508 AATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTA 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTA 422

Qy        568 AAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCC 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        423 AAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCC 482

Qy        628 CAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGA 542

Qy        688 ATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAA 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 ATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAA 602

Qy        748 TGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAAT 807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 TGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAAT 662

Qy        808 GCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCA 867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 GCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCA 722

Qy        868 AGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCC 927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 AGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCC 782

Qy        928 CCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTG 987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 CCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTG 842

Qy        988 GATGATGATTATCTTGGGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGAG 1047
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        843 GATGATGATTATCTTGAGAATGAAGAGGAATATGAAGATGTTGTAGAAGATGTAAAAGAG 902

Qy       1048 GAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGA 1107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGA 962

Qy       1108 GGACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATTACTG 1167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GGACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAGAAATTACTG 1022

Qy       1168 ACTGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGCCTAC 1227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 ACTGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGCCTAC 1082

Qy       1228 AAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||||||||||||||||
Db       1083 AAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1115
```

Search completed: November 29, 2004, 14:49:20
Job time : 5514 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:25 ; Search time 653.5 Seconds
(without alignments)
10121.301 Million cell updates/sec

Title: US-10-798-532-3
Perfect score: 1260
Sequence: 1 atggcgggatccggctgcgc.....tgtatcaaaccttgctttaa 1260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1260	100.0	1260	3	AAZ51356	Aaz51356 Human tra
	2	1258.4	99.9	1260	3	AAZ51355	Aaz51355 Human wil
	3	1256.8	99.7	2008	6	ABQ99383	Abq99383 Human cod
	4	1255.2	99.6	1260	4	AAI66829	Aai66829 Human pro
	5	1255.2	99.6	2622	4	AAI66828	Aai66828 Human pro
	6	1255.2	99.6	4065	4	AAH72714	Aah72714 Human cer
	7	1255.2	99.6	4065	5	ADL45238	Adl45238 Human ova
	8	1253.6	99.5	1260	4	AAS06734	Aas06734 Polynucle
	9	1253.6	99.5	1824	4	AAD18818	Aad18818 Human kin
	10	1173.4	93.1	3998	12	ADQ24401	Adq24401 Human sof
	11	1106.4	87.8	3244	10	ADB52857	Adb52857 Primary r
	12	458.6	36.4	529	4	AAH70764	Aah70764 Human cer
	13	445.8	35.4	722	3	AAA02535	Aaa02535 Human col
	14	426.6	33.9	470	4	AAH71966	Aah71966 Human cer
	15	375.4	29.8	461	4	AAH69101	Aah69101 Human cer
	16	272.2	21.6	411	6	ABQ99225	Abq99225 Human ORF
	17	203.8	16.2	231	3	AAA01259	Aaa01259 Human col
	18	177.4	14.1	419	5	ADI73740	Adi73740 Human ova
	19	177.4	14.1	419	5	ADI67345	Adi67345 Human ova
c	20	176.2	14.0	546	6	ABQ53127	Abq53127 Oligonucl
	21	176.2	14.0	546	6	ABQ53126	Abq53126 Oligonucl
c	22	176	14.0	547	6	ABQ44465	Abq44465 Oligonucl
	23	176	14.0	547	6	ABQ44464	Abq44464 Oligonucl
	24	157.2	12.5	546	6	ABQ53129	Abq53129 Oligonucl
c	25	157.2	12.5	546	6	ABQ53128	Abq53128 Oligonucl
	26	157.2	12.5	547	6	ABQ44463	Abq44463 Oligonucl
c	27	157.2	12.5	547	6	ABQ44462	Abq44462 Oligonucl
	28	148.4	11.8	442	9	ACH28557	Ach28557 Human adu
	29	85	6.7	198	5	ADL38977	Adl38977 Human ova
	30	63.4	5.0	65	6	ABN52650	Abn52650 Mouse spl
	31	63.2	5.0	549	10	ADK58517	Adk58517 Plant DNA
	32	57.8	4.6	1196	12	ADJ39614	Adj39614 Plant cDN
	33	53.8	4.3	1364	12	ADJ42189	Adj42189 Plant cDN
	34	47.4	3.8	1747	6	AAS94774	Aas94774 Human DNA
	35	47.4	3.8	1863	3	AAZ47544	Aaz47544 Human pro
	36	47.4	3.8	2044	8	ABX70932	Abx70932 Novel hum
	37	47.4	3.8	2887	12	ADJ96566	Adj96566 Human DYR
	38	46.8	3.7	7565	12	ADK70703	Adk70703 Murine CR
	39	45.4	3.6	1581	6	ABK88905	Abk88905 Murine mD
	40	44.4	3.5	2658	3	AAA70213	Aaa70213 Plasmodiu
	41	43.8	3.5	1584	6	ABL58336	Abl58336 Rat YAK1
	42	43	3.4	681	10	ADK52877	Adk52877 Plant DNA
c	43	42.8	3.4	600	12	ADJ44008	Adj44008 Plant cDN
	44	42.4	3.4	1497	3	AAA70154	Aaa70154 Plasmodiu
	45	42.2	3.3	2095	3	AAC36941	Aac36941 Arabidops

ALIGNMENTS

RESULT 1
AAZ51356
ID AAZ51356 standard; DNA; 1260 BP.
XX
AC AAZ51356;
XX
DT 06-JUN-2000 (first entry)
XX
DE Human transdominant mutant serine/threonine kinase KIS (hKIS) gene.
XX
KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
KW transdominant gene; mutant; cyclin-dependent kinase inhibitor; CKI; p27;
KW modulator; treatment; cell proliferative disease; vascular disorder;
KW gene therapy; atherosclerosis; restenosis; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1260
FT /*tag= a
FT /product= "Human transdominant mutant KIS protein"
FT /transl_except= (pos:556. .561, Asn)
FT misc_feature 161
FT /*tag= b
FT /note= "A to G substitution resulting in a Lys to Arg
FT change in the amino acid sequence"
XX
PN WO200011165-A1.
XX
PD 02-MAR-2000.
XX
PF 20-AUG-1999; 99WO-US018903.
XX
PR 21-AUG-1998; 98US-0097710P.
XX
PA (NABE/) NABEL G J.
PA (NABE/) NABEL E G.
XX
PI Nabel GJ, Nabel EG;
XX
DR WPI; 2000-237648/20.
DR P-PSDB; AAY70306.
XX
PT Novel serine/threonine kinase hKIS polynucleotides and polypeptides used

PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT proliferation.
XX
PS Claim 7; Page 59; 70pp; English.
XX
CC The present DNA sequence encodes a transdominant mutant human KIS (hKIS),
CC constructed by site directed mutagenesis. A single nucleotide
CC substitution (A to G) results in a lysine to arginine change in the
CC protein sequence. hKIS is a serine/threonine kinase, that acts as an
CC inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
CC controls cell proliferation and is localised predominantly in the
CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC G1 phase. The hKIS sequences are used to modulate cell proliferation and
CC treat cell proliferative and vascular diseases. The polynucleotide
CC sequence may be used in gene therapy to treat vascular disorders such as
CC restenosis or atherosclerosis
XX
SQ Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;

Query Match 100.0%; Score 1260; DB 3; Length 1260;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

RESULT 2
AAZ51355
ID AAZ51355 standard; DNA; 1260 BP.
XX
AC AAZ51355;
XX
DT 06-JUN-2000 (first entry)
XX
DE Human wild type serine/threonine kinase KIS (hKIS) gene.
XX
KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
KW dominant gene; cyclin-dependent kinase inhibitor; CKI; p27; modulator;
KW treatment; cell proliferative disease; vascular disorder; gene therapy;
KW atherosclerosis; restenosis; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1260
FT /*tag= a
FT /product= "Human wild type KIS protein"
FT /transl_except= (pos:556. .561, Asn)
XX
PN WO200011165-A1.
XX
PD 02-MAR-2000.
XX
PF 20-AUG-1999; 99WO-US018903.
XX
PR 21-AUG-1998; 98US-0097710P.
XX
PA (NABE/) NABEL G J.
PA (NABE/) NABEL E G.
XX
PI Nabel GJ, Nabel EG;
XX
DR WPI; 2000-237648/20.
DR P-PSDB; AAY70305.

XX
PT Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT proliferation.
XX
PS Claim 1; Page 56; 70pp; English.
XX
CC The present DNA sequence encodes a wild type human KIS (hKIS), isolated
CC from a human B-cell library, using a yeast two hybrid screening system.
CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
CC that controls cell proliferation and is localised predominantly in the
CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
CC The hKIS sequences are used to modulate cell proliferation and treat cell
CC proliferative and vascular diseases. The polynucleotide sequence may be
CC used in gene therapy to treat vascular disorders such as restenosis or
CC atherosclerosis
XX
SQ Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;

Query Match 99.9%; Score 1258.4; DB 3; Length 1260;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

RESULT 3
ABQ99383
ID ABQ99383 standard; cDNA; 2008 BP.
XX
AC ABQ99383;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human coding sequence SEQ ID 116.
XX
KW Human; expressed sequence tag; EST; haematopoietic disorder;
KW central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200259260-A2.
XX
PD 01-AUG-2002.
XX
PF 16-NOV-2001; 2001WO-US042950.
XX
PR 17-NOV-2000; 2000US-00714936.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-590824/63.
DR N-PSDB; ABP64797.

```
XX
PT   New isolated polynucleotide, useful in research, diagnostic or
PT   therapeutic methods, e.g. preventing or treating disorders involving
PT   aberrant protein expression or biological activity.
XX
PS   Claim 1; SEQ ID NO 116; 394pp; English.
XX
CC   The present invention relates to novel human coding sequences (ABQ99268-
CC   ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC   therapeutic, diagnostic and research methods. The polynucleotides may be
CC   used in the field of molecular biology as hybridisation probes, primers
CC   for PCR, for chromosome and gene mapping, for the recombinant production
CC   of protein, or in generation of anti-sense DNA or RNA. The
CC   polynucleotides are useful in diagnostics as expressed sequence tags
CC   (ESTs) for identifying expressed genes or for physical mapping of the
CC   human genome. The proteins may be used as molecular weight markers, or as
CC   nutritional sources or supplements. The proteins may be used to maintain
CC   and expand cell population in a totipotential or pluripotential state
CC   useful for re-engineering damaged or diseased tissues, transplantation,
CC   manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC   polynucleotides and proteins are useful for preventing, treating or
CC   ameliorating disorders involving aberrant protein expression or
CC   biological activity, e.g. haematopoietic disorders, central/peripheral
CC   nervous system diseases, mechanical and traumatic disorders, non-healing
CC   wounds, immune deficiencies and disorders, infectious diseases caused by
CC   viral, bacterial or fungal infection, autoimmune disorders, allergic
CC   reactions and conditions, coagulation disorders, or cancer. The
CC   polynucleotide sequences of the invention were assembled from ESTs
CC   isolated mainly by sequencing by hybridisation, and in some cases,
CC   sequences obtained from one or more public databases. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;

  Query Match             99.7%;  Score 1256.8;  DB 6;  Length 2008;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1258;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 138

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 198

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        199 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 258

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 318

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 378

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 438

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 498

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 558

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 618

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 678

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 738

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 798

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 858

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 918

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 978

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        979 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1038

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1039 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1098

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1099 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1158

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1159 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1218

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1219 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1278

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1279 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1338

RESULT 4
AAI66829
ID   AAI66829 standard; cDNA; 1260 BP.
XX
AC   AAI66829;
XX
DT   07-JAN-2002  (first entry)
XX
DE   Human protein kinase polypeptide 13887 coding sequence.
XX
KW   Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KW   cytostatic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1260
FT                   /*tag=  a
FT                   /product= "protein kinase 13887"
XX
```

PN WO200173050-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-US009483.
XX
PR 24-MAR-2000; 2000US-0191846P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers R;
XX
DR WPI; 2001-611632/70.
DR P-PSDB; AAG65767.
XX
PT New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT useful in diagnosis of cancer or cellular proliferation or
PT differentiation disorders and to screen for polypeptide modulators useful
PT to treat such conditions.
XX
PS Claim 1; Fig 10A-C; 169pp; English.
XX
CC The invention provides novel human protein kinase polypeptides, 3714,
CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC protein kinase polypeptides can be expressed by standard recombinant
CC methodology.3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC are useful for diagnostic and screening methods to identify subjects (at
CC risk of) having cancer or cellular proliferation and/or differentiation
CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC modulators are useful for the treatment of cancer, particularly colon
CC cancer or cellular proliferation and/or differentiation disorders. Other
CC disorders associated with 3714, 16742, 23546 or 13887 expression or
CC activity that can be treated include bone related disorders, inflammatory
CC disorders, autoimmune diseases, cardiovascular disorders and liver
CC diseases. The present sequence represents a human protein kinase
CC polypeptide 13887 coding sequence
XX
SQ Sequence 1260 BP; 312 A; 279 C; 334 G; 335 T; 0 U; 0 Other;

Query Match 99.6%; Score 1255.2; DB 4; Length 1260;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

RESULT 5
AAI66828
ID AAI66828 standard; cDNA; 2622 BP.
XX
AC AAI66828;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human protein kinase polypeptide 13887 encoding cDNA.
XX
KW Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KW cytostatic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1. .266

```
FT                  /*tag=  a
FT   CDS            267. .1529
FT                  /*tag=  b
FT                  /product= "protein kinase 13887"
FT                  /note= "coding sequence specifically claimed"
FT   3'UTR          1530. .2622
FT                  /*tag=  c
XX
PN   WO200173050-A2.
XX
PD   04-OCT-2001.
XX
PF   23-MAR-2001; 2001WO-US009483.
XX
PR   24-MAR-2000; 2000US-0191846P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Meyers R;
XX
DR   WPI; 2001-611632/70.
DR   P-PSDB; AAG65767.
XX
PT   New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT   useful in diagnosis of cancer or cellular proliferation or
PT   differentiation disorders and to screen for polypeptide modulators useful
PT   to treat such conditions.
XX
PS   Claim 1; Fig 10A-C; 169pp; English.
XX
CC   The invention provides novel human protein kinase polypeptides, 3714,
CC   16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC   protein kinase polypeptides can be expressed by standard recombinant
CC   methodology.3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC   are useful for diagnostic and screening methods to identify subjects (at
CC   risk of) having cancer or cellular proliferation and/or differentiation
CC   disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC   modulators are useful for the treatment of cancer, particularly colon
CC   cancer or cellular proliferation and/or differentiation disorders. Other
CC   disorders associated with 3714, 16742, 23546 or 13887 expression or
CC   activity that can be treated include bone related disorders, inflammatory
CC   disorders, autoimmune diseases, cardiovascular disorders and liver
CC   diseases. The present sequence represents a human protein kinase
CC   polypeptide 13887 encoding cDNA
XX
SQ   Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other;

  Query Match           99.6%;  Score 1255.2;  DB 4;  Length 2622;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1257;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 326

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 386

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        387 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 446

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 506

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 566

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        567 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 626

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        627 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 686

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 746

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 806

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 866

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 926

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 986

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1046

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1106

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1166

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1226

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1227 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1286

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1346

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1406

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1466

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1526


RESULT 6
AAH72714
ID   AAH72714 standard; cDNA; 4065 BP.
XX
AC   AAH72714;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 3988.
XX
```

```
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer and
PT   for assessing and detecting compounds for treating the cancer.
XX
PS   Claim 1; Page 810-811; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

  Query Match             99.6%;  Score 1255.2;  DB 4;  Length 4065;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1257;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1527


RESULT 7
ADL45238
ID   ADL45238 standard; DNA; 4065 BP.
XX
AC   ADL45238;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human ovarian cancer DNA marker #19128.
XX
KW   Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
OS   Homo sapiens.
XX
PN   WO200170979-A2.
XX
PD   27-SEP-2001.
XX
```

PF 21-MAR-2001; 2001WO-US009126.
XX
PR 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lee J, Lillie J;
XX
DR WPI; 2001-611502/70.
XX
PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX
PS Disclosure; SEQ ID NO 19128; 106pp; English.
XX
CC The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.
XX
SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

Query Match 99.6%; Score 1255.2; DB 5; Length 4065;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1527
```

RESULT 8
AAS06734

```
ID   AAS06734 standard; cDNA; 1260 BP.
XX
AC   AAS06734;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Polynucleotide sequence encoding human protein kinase #34.
XX
KW   Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW   metabolic disorder; immune related disease; neurological disorder;
KW   neurodegenerative disorder; inflammatory disorder; infectious disease;
KW   reproductive disorder; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200138503-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US032085.
XX
PR   24-NOV-1999;   99US-0167482P.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman GD,  Whyte D,  Manning G,  Sudarsanam S,  Martinez R;
PI   Flanagan P,  Clary D;
XX
DR   WPI; 2001-343950/36.
DR   P-PSDB; AAU03534.
XX
PT   Nucleic acids encoding human kinase polypeptides, useful for preventing
PT   diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT   neuronal-associated diseases, and microbial infections.
XX
PS   Example 1; Fig 1; 433pp; English.
XX
CC   AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
CC   protein kinases have been identified as members of the tyrosine or
CC   serine/threonine kinase (PTK and STK) families. The polynucleotides
CC   encoding protein kinases and the polypeptides may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate kinase expression. For example, they may be used to treat
CC   cancers (especially cancers of haematopoietic origin), cardiovascular
CC   disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC   immune related diseases (e.g. rheumatoid arthritis), neurological
CC   disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC   Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC   disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC   Additionally, polynucleotides encoding protein kinases may be used for
CC   gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC   polypeptides may be used as antigens in the production of antibodies
CC   against the protein kinases and in assays to identify modulators of
CC   protein kinase expression and activity
XX
SQ   Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other;

  Query Match             99.5%;  Score 1253.6;  DB 4;  Length 1260;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1256;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
```

```
RESULT 9
AAD18818
ID   AAD18818 standard; cDNA; 1824 BP.
XX
AC   AAD18818;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human kinase (PKIN)-3 cDNA.
XX
KW   Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
KW   cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
KW   AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
KW   atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
KW   myasthenia gravis; cirrhosis; cataract; growth and development disorder;
KW   seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW   lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW   asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
KW   antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             169. .1428
FT                   /*tag= a
FT                   /product= "Human PKIN-3 protein"
XX
PN   WO200181555-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-US012992.
XX
PR   20-APR-2000; 2000US-0199021P.
PR   28-APR-2000; 2000US-0200226P.
PR   05-MAY-2000; 2000US-0202339P.
PR   11-MAY-2000; 2000US-0203505P.
PR   18-MAY-2000; 2000US-0205564P.
PR   26-MAY-2000; 2000US-0207739P.
PR   01-JUN-2000; 2000US-0208795P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
PI   Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Walia NK, Yao MG;
PI   Patterson C, Burrill JD, Marcus GA, Zingler KA, Recipon SA, Lu Y;
PI   Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
PI   Walsh RT, Ramkumar J, Borowsky ML, Au-Young J, Hillman JL;
PI   Gururajan R;
XX
DR   WPI; 2001-611740/70.
DR   P-PSDB; AAE11769.
XX
PT   Human kinases and nucleic acids, useful for preventing diagnosing and
PT   treating cancers, inflammation and immune disorders.
XX
PS   Claim 5; Page 152; 166pp; English.
XX
CC   The present invention relates to human kinases (PKIN) and the nucleic
CC   acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
CC   used in the prevention, diagnosis and treatment of diseases cancers,
CC   adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
CC   acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
CC   gout, microbial infections, cardiovascular disease and/or inflammation,
CC   myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
CC   infarction, cataract, growth and development disorder, seizure disorder,
CC   pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
CC   disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
CC   PKIN may be used to treat disorders associated with decreased PKIN
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of PKIN by expressing inactive proteins or to
CC   supplement the patients own production of PKIN. PKIN nucleic acids may be
CC   used to produce the PKIN polypeptide, by inserting the nucleic acids into
CC   a host cell and culturing the cell to express the protein. PKIN nucleic
CC   acid and its complementary sequences may also be used as DNA probes in
CC   diagnostic assays to detect and quantitate the presence of similar
CC   nucleic acid sequences in samples and therefore which patients may be in
CC   need of restorative therapy. The present sequence is human PKIN-3 cDNA
XX
SQ   Sequence 1824 BP; 434 A; 416 C; 472 G; 502 T; 0 U; 0 Other;

  Query Match           99.5%;  Score 1253.6;  DB 4;  Length 1824;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 1256;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 228

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 288

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        289 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 348

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 408

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 468

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 528

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        529 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 588

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 648

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 708

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 768

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 828

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 888

Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 948

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1008

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1068

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1128

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
                ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1129 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1188

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1189 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1248

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
                ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1249 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1308

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1309 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1368

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1369 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1428
```

RESULT 10
ADQ24401
ID ADQ24401 standard; DNA; 3998 BP.
XX
AC ADQ24401;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7221.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 7221; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 3998 BP; 1066 A; 777 C; 885 G; 1255 T; 0 U; 15 Other;

  Query Match   93.1%; Score 1173.4; DB 12; Length 3998;
  Best Local Similarity 96.1%; Pred. No. 0;
  Matches 1214; Conservative   0; Mismatches   46; Indels   3; Gaps   1;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 219

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 279

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        280 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 339

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              |||||||||||||||||||||||||||||||||||||||| |||          |   | |
Db        340 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTAAGCATGTTTCTTACCAGACCG 399

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              | |   |  | | |   | ||||||||||||||||||||||||||||||  ||
Db        400 AAGGTCTGTGTTGACCTTAACCGGAGAGTGACTTTGTATGGAGTGTTTACAATCCAGGTT 459

Qy        301 TCT---CCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCG 357
                |   | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 GTTCTCCAAAGTGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCG 519

Qy        358 GAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCC 417
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        520 GAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCC 579

Qy        418 CGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTC 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 CGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTC 639

Qy        478 AAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGA 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 AAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGA 699

Qy        538 CTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCT 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 CTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCT 759

Qy        598 CCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGT 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 CCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGT 819

Qy        658 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 879

Qy        718 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 939

Qy        778 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 999

Qy        838 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 1059

Qy        898 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCC 1119

Qy        958 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAA 1017
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
```

```
Db       1120 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAA 1179

Qy       1018 TATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCT 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 TATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCT 1239

Qy       1078 CTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCT 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 CTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCT 1299

Qy       1138 GGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 GGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTT 1359

Qy       1198 GTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1360 GTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1419

Qy       1258 TAA 1260
              |||
Db       1420 TAA 1422
```

RESULT 11
ADB52857
ID ADB52857 standard; DNA; 3244 BP.
XX
AC ADB52857;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3399.
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN WO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
DR WPI; 2003-731472/69.
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 3399; 874pp; English.
XX
CC The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 3244 BP; 800 A; 710 C; 819 G; 915 T; 0 U; 0 Other;

Query Match 87.8%; Score 1106.4; DB 10; Length 3244;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 1164; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
Qy          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60
              ||||| || |||||||||||| |||||||| |||||||||| || ||||||||||||||||
Db        162 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 221

Qy         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120
              ||||||||||||||||||||||| || || |||||||||| |||||||||| |||||
Db        222 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 281

Qy        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180
              |||||||||||||  || |||||||||||||||||||||| ||||||| |||| || |||
Db        282 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 341

Qy        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240
              ||||||||||||||  ||||| || || ||||||||||||||| |||||||||| ||||||
Db        342 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 401

Qy        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300
              || ||||||||||||||| |||||||||||||| || ||||| ||||| || |||||
Db        402 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 461

Qy        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        462 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 521

Qy        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420
              |||||  |||||||||||||||||||| |||||||||||||| ||||| || ||| ||
Db        522 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 581

Qy        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480
              |||||  |||||||||||||||||||| |||||||||||||||||||| ||||||||
Db        582 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 641

Qy        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540
              ||||| |||||  |||||||||| ||||| || || ||||| || ||||||||||||||
Db        642 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 701

Qy        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
              |||||||||||||||||||||| || ||||||||||||||||||||||||||| |||||
Db        702 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 761

Qy        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660
              ||||||||| |||| |||||||||||||||| |||||||||||||||||||| ||||||
Db        762 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 821

Qy        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720
              |||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db        822 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 881
```

```
Qy        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        882 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 941

Qy        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        942 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 1001

Qy        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900
              |||||||||||||||||| || |||||||||| || |||||||| ||||| ||||||
Db       1002 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 1061

Qy        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||
Db       1062 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 1121

Qy        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020
              |||||||| ||||| || |||||||||||||||| |||||||  |||||||| ||||||
Db       1122 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1181

Qy       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080
              |||||||||||||||||||||||||| |||||||| |||||||||||||||||| |||||
Db       1182 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1241

Qy       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||||
Db       1242 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1301

Qy       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200
              |||||||||||||| |||||||| |||||||| |||||||||| ||||||||||||||||
Db       1302 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1361

Qy       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTTTAA 1260
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       1362 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTTTAA 1421
```

RESULT 12
AAH70764
ID AAH70764 standard; cDNA; 529 BP.
XX
AC AAH70764;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 2038.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 434; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

Query Match 36.4%; Score 458.6; DB 4; Length 529;
Best Local Similarity 97.8%; Pred. No. 7.1e-126;
Matches 486; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

```
Qy        654 ATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 713
              | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 AGGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 81

Qy        714 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 773
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 141

Qy        774 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 833
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 201

Qy        834 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 893
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 261

Qy        894 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 953
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 321

Qy        954 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGA 1013
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        322 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGA 381

Qy       1014 GGAATATGAAGA-TGTTGTAGAAGATGTAAAAGA-GGAGTGTCAAAAATATGGACCAGTG 1071
              |||||||||||| |||||||||||||||||||| |||||||||||||||||||||||| |
Db        382 GGAATATGAAGATTGTTGTAGAAGATGTAAAAGAGGGAGTGTCAAAAATATGGACCAGGG 441

Qy       1072 GTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCA 1131
              |||||||||||| |||||||||||||||||||||  ||||||||||||||||||||
Db        442 GTATCTCTACTTGGTCCAAAGGAAAATCCTGGCAGAGGAACAGTCTTTGTTGAGTATGCA 501

Qy       1132 AATGCTGGTGATTCCAA 1148
              || ||||| ||||| ||
Db        502 AAGGCTGGGGATTCAAA 518
```

RESULT 13
AAA02535
ID AAA02535 standard; cDNA; 722 BP.
XX
AC AAA02535;
XX
DT 19-MAY-2000 (first entry)
XX
DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:2526.
XX
KW Human; colon cancer; tumour; diagnosis; gene expression product; probe;
KW detection; cancerous state; metastasis; identification; breast cancer;
KW oestrogen receptor-positive breast cancer; therapy;
KW oestrogen receptor-negative breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
PN WO9958675-A2.

```
XX
PD   18-NOV-1999.
XX
PF   13-MAY-1999;   99WO-US010602.
XX
PR   14-MAY-1998;   98US-0085426P.
PR   15-MAY-1998;   98US-0085537P.
PR   15-MAY-1998;   98US-0085696P.
PR   21-OCT-1998;   98US-0105234P.
PR   27-OCT-1998;   98US-0105877P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI   Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI   Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI   Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
DR   WPI; 2000-126369/11.
XX
PT   Polynucleotide library used to determine cancerous states of mammalian
PT   cells.
XX
PS   Claim 1; Page 1020; 1097pp; English.
XX
CC   AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC   libraries constructed from human colon cancer cell lines. The present
CC   invention also describes a method of detecting differentially expressed
CC   genes correlated with a cancerous state of a mammalian cell, comprising
CC   detecting at least one differentially expressed gene product in a test
CC   sample derived from a cell suspected of being cancerous, where detection
CC   of the differentially expressed gene product is correlated with a
CC   cancerous state of the cell from which the test sample was derived. The
CC   polynucleotides sequences can be used in a method for detecting
CC   differentially expressed genes correlated with a cancerous state of a
CC   mammalian cell. The polynucleotides can also be used as probes for
CC   detecting and mapping related genes. They can be used in diagnosis and
CC   prognosis of diseases and disorders (e.g. identification of pre-
CC   metastatic or metastatic cancerous states, stages of cancer, or
CC   responsiveness of cancer to therapy). This is particularly for breast
CC   cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC   negative breast cancer, lung cancer, and colon cancer
XX
SQ   Sequence 722 BP; 215 A; 129 C; 164 G; 193 T; 0 U; 21 Other;

  Query Match             35.4%;  Score 445.8;  DB 3;  Length 722;
  Best Local Similarity   93.0%;  Pred. No. 5.6e-122;
  Matches  581;  Conservative    0;  Mismatches   29;  Indels   15;  Gaps   11;

Qy        454 GAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAA 513
              |||||||||||||||||| |||||||| |||||||||||||||||||||||||||||||
Db          1 GAGGGCTATGTCCATGCGGNCCTCAAACNACGTAACATATTGTGGAGTGCAGAGAATGAA 60

Qy        514 TGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 573
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db         61 TGTTTTAAACTCATTGACTTTGGACTTANCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 120

Qy        574 ATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCT 633
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||| |||
Db        121 ATTCAGACAGACGGGTATCGGGCTCCAGAA-CAGAATTGCAAAATTGCTTGGCCCANGCT 179

Qy        634 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 693
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 239

Qy        694 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 753
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 299

Qy        754 GCAAACAG-TTCTGCTATTA-TTGATCACATATTTGCCAGTAAA--GCAGTGGTGAATGC 809
              |||||||| |||||||||||| |||||||||||||||||||||    | ||||||||||
Db        300 GCAAACAGTTTCTGCTATTATTTGATCACATATTTGCCAGTAAAAGCAANTGGTGAATGC 359

Qy        810 CGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCC--A 867
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||  |
Db        360 CGCAATTCCAGCCTATCACCTAANAGACCTTATCAAAAGCATGCTTCATGATGATCCCAA 419

Qy        868 AGCAGAAGAATTCCTGCTGAAATGGCATT-GTGCAGCCCATTCTTTAGC-ATTCCTTTTG 925
              |||||||| || |||||||||||||| ||||| |||||||||||||| |||||||||
Db        420 GCAGGAAGAATTTCTNCTGAAATGGCATTGGTGCANCCCATTCTTTAGCNATTCCTTTTG 479

Qy        926 CCCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGT-GCTAAGACTGCTGAATGTG 984
              |||||||||||||||||| |||||||| |||||||||| ||||||||||||||||||||
Db        480 CCCCTCATATTGAAGATCTGNTCATGCTTTCCACTCCAGTGGCTAAGACTGCTGAATGTG 539

Qy        985 CT-GGATGATGATTATCTTGGGAATGAA--GAGGAATATGAAGATGTTGTAGAAGATGT- 1040
              || || |||||||||||||| |||||||    |||| |||||||||||||| |||||||
Db        540 CTGGGNTGATGATTATCTTGAGAATGAAAGAAGGATTATGAAGATGTTGTTGAAGATGNT 599

Qy       1041 -AAAAGAGGAGTGTCAAAAATATGG 1064
               |||  || |||| ||||||| | |
Db        600 AAAAGAAGAAGTGGCAAAAATNTTG 624


RESULT 14
AAH71966
ID   AAH71966 standard; cDNA; 470 BP.
XX
AC   AAH71966;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 3240.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer and
PT   for assessing and detecting compounds for treating the cancer.
XX
PS   Claim 1; Page 636; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;

  Query Match             33.9%;  Score 426.6;  DB 4;  Length 470;
  Best Local Similarity   98.7%;  Pred. No. 2.4e-116;
```

```
  Matches  451;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy        654 ATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 713
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         14 AGGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 73

Qy        714 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 133

Qy        774 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 193

Qy        834 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 253

Qy        894 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 313

Qy        954 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGA 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        314 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGA 373

Qy       1014 GGAATATGAAGA-TGTTGTAGAAGATGTAAAAGA-GGAGTGTCAAAAATATGGACCAGTG 1071
              |||||||||||| |||||||||||||||||||| |||||||||||||||||||||||| |
Db        374 GGAATATGAAGATTGTTGTAGAAGATGTAAAAGAGGGAGTGTCAAAAATATGGACCAGGG 433

Qy       1072 GTATCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAG 1108
              |||||||||||||| ||||||||||||||||||||||
Db        434 GTATCTCTACTTGGTCCAAAGGAAAATCCTGGCAGAG 470


RESULT 15
AAH69101
ID   AAH69101 standard; cDNA; 461 BP.
XX
AC   AAH69101;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 375.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer and
PT   for assessing and detecting compounds for treating the cancer.
XX
PS   Claim 1; Page 168-169; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 461 BP; 134 A; 91 C; 109 G; 127 T; 0 U; 0 Other;

  Query Match             29.8%;  Score 375.4;  DB 4;  Length 461;
  Best Local Similarity   96.2%;  Pred. No. 4.4e-101;
  Matches  428;  Conservative    0;  Mismatches   11;  Indels    6;  Gaps    4;

Qy        654 ATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 713
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 AGGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGG 77

Qy        714 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 AATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTAT 137

Qy        774 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 TGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAG 197

Qy        834 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGC 257

Qy        894 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCT 953
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        258 ATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATA-TGAAGATCTGGTCATGCT 316

Qy        954 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGA 1013
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        317 TCCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGA 376

Qy       1014 GGAATATGAAGAT--GTTGTAGAAGATGTAAAAGAGG--AGTGTCAAAAATATGGACCA- 1068
              |||||||||||||    | ||||||||||||||||    |||||||||||||||||| |
Db        377 GGAATATGAAGATTGTTGTTAGAAGATGTAAAAGAGGGAGGTGTCAAAAATATGGACCAG 436

Qy       1069 GTGGTATCTCTACTTGTTCCAAAGG 1093
              |||||||||||||| | | || |
Db        437 GTGGTATCTCTACTTTGTTCCAAAG 461


Search completed: November 29, 2004, 11:45:27
Job time : 656.5 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 3159.5 Seconds
(without alignments)
4832.489 Million cell updates/sec

Title: US-10-798-532-4
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAFG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10798532/runat_29112004_112453_15901/app_query.fasta_1.1166
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532_@CGN_1_1_6628_@runat_29112004_112453_15901 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2171	98.4	1763	3	AK013347	AK013347 Mus muscu
2	1711	77.6	2846	3	AK030152	AK030152 Mus muscu
3	1521.5	69.0	1079	4	BM468107	BM468107 AGENCOURT
4	1271.5	57.6	1015	6	BY712344	BY712344 BY712344
5	1242	56.3	917	4	BM451184	BM451184 AGENCOURT
6	1169	53.0	823	7	CO771206	CO771206 testis_ES
7	1162	52.7	779	5	BU447831	603767571
8	1158	52.5	896	7	CO773091	CO773091 testis_ES
9	1133	51.4	781	7	CN232453	CN232453 WLB074D11
10	1107	50.2	932	5	BU541306	BU541306 AGENCOURT
11	1106	50.1	717	5	BU385761	BU385761 603582650
12	1091	49.5	899	2	BE895119	BE895119 601436041
13	1079	48.9	772	4	BI562420	BI562420 603256316
14	1075	48.7	794	5	BU205357	BU205357 603104569
15	1070.5	48.5	834	7	CF593363	CF593363 AGENCOURT
16	1068.5	48.4	875	4	BI259568	BI259568 602968330
17	1063	48.2	888	5	BQ225071	BQ225071 AGENCOURT
18	1050	47.6	657	4	BG547675	BG547675 602575520
19	1050	47.6	682	4	BI333581	BI333581 602997069
20	1025	46.5	844	4	BG474616	BG474616 602517423
21	1014	46.0	602	7	CO737403	CO737403 S1LT02c17
22	1005.5	45.6	876	5	BU184346	BU184346 AGENCOURT
23	990.5	44.9	906	4	BG386905	BG386905 602454704
24	968.5	43.9	696	5	BU439278	BU439278 604147528
25	967	43.8	631	5	BU112717	BU112717 603131421
26	966.5	43.8	1003	5	BQ922020	BQ922020 AGENCOURT
27	962	43.6	755	5	BU371858	BU371858 603568233
28	955	43.3	699	5	BU707284	BU707284 UI-M-FR0-
29	944	42.8	604	5	BX261107	BX261107 BX261107
30	909.5	41.2	716	7	CF745774	CF745774 UI-M-GV0-
31	900	40.8	570	5	BU293329	BU293329 604166812
32	893	40.5	670	7	CN359673	CN359673 170004245
33	889	40.3	928	5	BQ424438	BQ424438 AGENCOURT
34	880.5	39.9	824	2	BE568528	BE568528 601342425
35	870.5	39.5	820	7	CK597557	CK597557 AGENCOURT
36	862	39.1	672	5	BX261106	BX261106 BX261106
37	858.5	38.9	1063	6	CD510246	CD510246 CDA100-G0
38	855	38.8	785	6	CB990192	CB990192 AGENCOURT
c 39	850	38.5	523	2	BF929410	BF929410 IL2-NT020
40	828.5	37.6	1063	5	BU356643	BU356643 603474124
41	818.5	37.1	743	4	BI562440	BI562440 603256324
42	815	36.9	763	7	CK364431	CK364431 AGENCOURT
43	798	36.2	933	2	BE886040	BE886040 601507192
c 44	796	36.1	986	5	BU751496	BU751496 CH4#002_E
45	793	35.9	484	5	BX281211	BX281211 BX281211

ALIGNMENTS

RESULT 1
AK013347
LOCUS AK013347 1763 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810454O04 product:kinase interacting with leukemia-associated gene (stathmin), full insert sequence.
ACCESSION AK013347
VERSION AK013347.1 GI:12850651
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,

Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1763)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.
FEATURES Location/Qualifiers
source 1. .1763
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2810454O04"
/db_xref="taxon:10090"
/clone="2810454O04"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10, 11 days embryo"
CDS 159. .1418
/note="unnamed protein product; kinase interacting with leukemia-associated gene (stathmin) (MGD|MGI:1341908) putative"
/codon_start=1
/protein_id="BAB28802.1"
/db_xref="GI:12850652"
/translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALXQLQGHRNIVTLYGVFTIHFSP
NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
CTSAVDLWSLGITLLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

Alignment Scores:

Pred. No.:	6.09e-216	Length:	1763
Score:	2171.00	Matches:	412
Percent Similarity:	98.81%	Conservative:	2
Best Local Similarity:	98.33%	Mismatches:	5
Query Match:	98.41%	Indels:	0
DB:	3	Gaps:	0

US-10-798-532-4 (1-419) x AK013347 (1-1763)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 218

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 278

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              |||||||||||| ||||||||||||||||||||||||||:::||||||||||||||||||||
Db        279 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 338

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 398

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 GANCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 458

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 518

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 578

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 638

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 698

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 758

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 818

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        819 TCAGCTGTTGATCTGTGGAGCCTCGGAATCACTTTACTGGAAATGTTCTCAGGAATGAAG 878

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAT 938

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 998

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
                 ||||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||
Db        999 ATCAAAAGCATGCTTCATGATGACCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGC 1058

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACT 1118

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
                 ||||||||||||||||||||||||||||||||||||||||||||    ||||||:::||||||
Db       1119 CCAGTGCTCAGGCTCCTCAATGTGCTGGATGATGATTACCTTGAAAATGAAGATGAATAT 1178

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1179 GAAGATGTTGTAGAAGATGTGAAAGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1238

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1239 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1298

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1299 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1358

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1359 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 1415
```

RESULT 2
AK030152

```
LOCUS       AK030152               2846 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:4933401D07 product:kinase interacting with
            leukemia-associated gene (stathmin), full insert sequence.
ACCESSION   AK030152
VERSION     AK030152.1  GI:26326136
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 2846)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.jp/
            URL:http://fantom.gsc.riken.jp/.
FEATURES             Location/Qualifiers
     source          1. .2846
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4933401D07"
                     /db_xref="taxon:10090"
                     /clone="4933401D07"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             238. .1230
                     /note="unnamed protein product; kinase interacting with
                     leukemia-associated gene (stathmin) (MGD|MGI:1341908)
                     putative"
                     /codon_start=1
                     /protein_id="BAC26812.1"
                     /db_xref="GI:26326137"
                     /translation="MTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMW
                     MIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY
                     IQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQE
                     WKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALCSPFFSIP
                     FAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSLLVPKE
                     NPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL"
     polyA_signal    2823. .2828
                     /note="putative"
     polyA_site      2846
                     /note="putative"
ORIGIN

Alignment Scores:
Pred. No.:              1.87e-167       Length:         2846
Score:                  1711.00         Matches:        326
Percent Similarity:     99.39%          Conservative:   2
```

```
Best Local Similarity:  98.79%        Mismatches:    2
Query Match:            77.56%        Indels:        0
DB:                     3             Gaps:          0

US-10-798-532-4 (1-419) x AK030152 (1-2846)

Qy          90 ValThrLeuTyrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
               :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         238 ATGACTTTATATGGAGTCTTTACCATACACTTCTCTCCCAATGTGCCATCACGCTGTCTG 297

Qy         110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGlnGly 129
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         298 CTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTTTTATATTCCAGTCATCAGGGC 357

Qy         130 CysSerMetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeu 149
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         358 TGCTCCATGTGGATGATACAGCACTGTGCCCGAGATGTTCTGGAGGCCCTCGCTTTTCTT 417

Qy         150 HisHisGluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGlu 169
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         418 CACCATGAGGGCTACGTCCATGCAGACCTCAAACCACGAAACATCCTGTGGAGTGCGGAG 477

Qy         170 AsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspVal 189
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         478 AATGAGTGTTTTAAGCTTATTGACTTTGGACTCAGCTTCAAAGAAGGCAATCAGGACGTA 537

Qy         190 LysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAla 209
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         538 AAGTATATTCAGACAGACGGGTATCGCGCTCCTGAAGCAGAGCTGCAGAACTGCTTGGCC 597

Qy         210 GlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGly 229
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         598 CAGGCCGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTCGGA 657

Qy         230 IleIleLeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGlu 249
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         658 ATCATTTTACTGGAAATGTTCTCAGGAATGAAGCTGAAACATACAGTCAGATCTCAGGAA 717

Qy         250 TrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAlaValValAsn 269
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         718 TGGAAGGCAAACAGTTCTGCTATTATTGATCATATATTTGCCAGTAAAGCAGTGGTGAAT 777

Qy         270 AlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspPro 289
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         778 GCCGCAATTCCAGCCTATCACCTCAGAGACCTTATCAAAAGCATGCTTCATGATGACCCG 837

Qy         290 SerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAla 309
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         838 GGCAGAAGAATCCCTGCTGAGATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCC 897

Qy         310 ProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeuAsnValLeu 329
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         898 CCTCATATTGAAGATCTGGTGATGCTTCCGACTCCAGTGCTCAGGCTCCTCAATGTGCTG 957

Qy         330 AspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAspValLysGlu 349
               ||||||||||||||   ||||||:::||||||||||||||||||||||||||||||||||
Db         958 GATGATGATTACCTTGAAAATGAAGATGAATATGAAGATGTTGTAGAAGATGTGAAAGAG 1017

Qy         350 GluCysGlnLysTyrGlyProValValSerLeuLeuValProLysGluAsnProGlyArg 369
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1018 GAGTGTCAGAAATATGGACCAGTGGTTTCTCTGCTTGTTCCAAAGGAAAATCCTGGCAGA 1077

Qy         370 GlyGlnValPheValGluTyrAlaAsnAlaGlyAspSerLysAlaAlaGlnLysLeuLeu 389
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1078 GGACAAGTCTTCGTTGAGTACGCGAACGCTGGTGATTCCAAAGCTGCTCAGAAGTTGCTG 1137

Qy         390 ThrGlyArgMetPheAspGlyLysPheValValAlaThrPheTyrProLeuSerAlaTyr 409
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1138 ACTGGGAGGATGTTTGACGGGAAGTTTGTTGTGGCTACATTCTACCCGCTGAGTGCCTAC 1197

Qy         410 LysArgGlyTyrLeuTyrGlnThrLeuLeu 419
               |||||||||||||||||||||||||||||
Db        1198 AAGAGGGGATATCTTTATCAAACCTTGCTT 1227
```

```
RESULT 3
BM468107
LOCUS       BM468107                1079 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6432293 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5535653
            5', mRNA sequence.
ACCESSION   BM468107
VERSION     BM468107.1  GI:18517149
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1079)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12224  row: b  column: 06
            High quality sequence stop: 754.
FEATURES             Location/Qualifiers
     source          1. .1079
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5535653"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
ORIGIN

Alignment Scores:
Pred. No.:              3.05e-148       Length:         1079
Score:                  1521.50         Matches:        309
Percent Similarity:     92.86%          Conservative:   3
Best Local Similarity:  91.96%          Mismatches:     17
Query Match:            68.97%          Indels:         7
DB:                     4               Gaps:           3

US-10-798-532-4 (1-419) x BM468107 (1-1079)

Qy          90 ValThrLeuTyrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           2 GTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTG 61

Qy         110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGlnGly 129
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 TTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGT 121

Qy         130 CysSerMetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeu 149
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 TGTTCCATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTT 181

Qy         150 HisHisGluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGlu 169
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         182 CATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAG 241

Qy         170 AsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspVal 189
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 AATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTA 301

Qy         190 LysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAla 209
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCC 361

Qy        210 GlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGly 229
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGA 421

Qy        230 IleIleLeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGlu 249
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAA 481

Qy        250 TrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAlaValValAsn 269
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAAT 541

Qy        270 AlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspPro 289
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCA 601

Qy        290 SerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAla 309
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCC 661

Qy        310 ProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeuAsnValLeu 329
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 CCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTGAATGTGCTG 721

Qy        330 AspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAspValLysGlu 349
                ||||||||||||||||  ||||||||||||||||||||||||||||:::  ||||||
Db        722 GATGATGATTATCTTGAGAATGAAGAGGAATATGAAGATGTTGTAGAAGAATGTAAAGAG 781

Qy        350 GluCysGlnLysTyrGlyProValValSerLeuLeuValProLysGluAsnProGlyArg 369
                ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db        782 GAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGNTCCAAAGGAAAATCCTGGCAGA 841

Qy        370 GlyGlnVal-PheValGluTyrAlaAsnAla-GlyAspSerLysAlaAlaGlnLys-Leu 388
                ||||||||| |||   ||||||||||||||   ||||||||||||||||||||| ||| 
Db        842 GGACAAGTCTTTTGNTGAGTATGCAAATGCTGNGTGATTCAAAAGCTGCGCAGAAATTTA 901

Qy        389 LeuThrGlyArg---MetPheAspGlyLysPheValValAla-ThrPheTyrProLeu-- 406
                |||||||||               |||||||||   ||||||  |||      |||   
Db        902 CTGACTGGNAANGATGGTTTGATGGGGAAGTTTGNTGTNGCCTACAATCCTACCCGCCTG 961

Qy        407 -SerAlaTyrLysArgGly---TyrLeuTyrGlnThrLeu 418
                 ||||||   ::::::|||      ||||||      |||
Db        962 GAGTGCCTTACAGAAGGGGATTATTCTGTATTCAAACCTT 1001

RESULT 4
BY712344
LOCUS       BY712344                1015 bp    mRNA    linear   EST 17-DEC-2002
DEFINITION  BY712344 RIKEN full-length enriched, 10, 11 days embryo whole body
            Mus musculus cDNA clone 2810454O04 5', mRNA sequence.
ACCESSION   BY712344
VERSION     BY712344.1  GI:27123621
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1015)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
  MEDLINE   22354683
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
            Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
            Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
            Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
            Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
            Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .1015
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="2810454O04"
                     /tissue_type="whole body"
                     /dev_stage="10, 11 days embryo"
                     /clone_lib="RIKEN full-length enriched, 10, 11 days embryo
                     whole body"
ORIGIN

Alignment Scores:
Pred. No.:              4.24e-122       Length:         1015
Score:                  1271.50         Matches:        251
Percent Similarity:     89.51%          Conservative:   5
Best Local Similarity:  87.76%          Mismatches:     23
Query Match:            57.64%          Indels:         7
DB:                     6               Gaps:           4
```

```
US-10-798-532-4 (1-419) x BY712344 (1-1015)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 218

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 278

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||   ||||||||||||||||||||||||:::||||||||||||||||||||
Db        279 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 338

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 398

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 GANCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 458

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 518

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 578

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 638

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 698

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 758

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 818

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||     |||||||||   ||||||||||||||||||||||||||||||
Db        819 TCAGCTGTTGATCTGTCGGGGCTCGGAATCACTTTACTGGAGATGTTCTCAGGAATGAAG 878

Qy        241 LeuLys---HisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
              ||||||   |||   :::          |||:::|||   ||||||
Db        879 CTGAAAATACATCAGATCTCAGGA------TGGGAGGCA---AGCAGTCTGCATCATTCG 929

Qy        260 HisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAsp 279
              |||||||||         ||||||   :::                |||       |||
Db        930 CATATATTT---------GCAGTGAGCAGTGGTGGTNCGCAATTCAGCCATCACTCAGAC 980

Qy        280 LeuIleLysSerMetLeu 285
                    :::|||||||||
Db        981 GCTTATCAGAGCATGCTC 998

RESULT 5
BM451184
LOCUS       BM451184                 917 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6392719 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5495159
            5', mRNA sequence.
ACCESSION   BM451184
VERSION     BM451184.1  GI:18500224
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12121  row: j  column: 24
            High quality sequence stop: 718.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5495159"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_67"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb.   Library constructed by Life
                     Technologies."
ORIGIN

Alignment Scores:
Pred. No.:              4.47e-119       Length:         917
Score:                  1242.00         Matches:        240
Percent Similarity:     99.59%          Conservative:   0
Best Local Similarity:  99.59%          Mismatches:     1
Query Match:            56.30%          Indels:         0
DB:                     4               Gaps:           0

US-10-798-532-4 (1-419) x BM451184 (1-917)

Qy        179 GlyLeuSerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArg 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 GGGCTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGG 81

Qy        199 AlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlu 218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAA 141

Qy        219 CysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGly 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 TGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGA 201

Qy        239 MetLysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIle 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 ATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATT 261

Qy        259 AspHisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArg 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGA 321

Qy        279 AspLeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAla 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCA 381

Qy        299 LeuCysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeu 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 TTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTT 441

Qy        319 ProThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGlu 338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        442 CCCACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAG 501
```

```
Qy        339 GluTyrGluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValVal 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 GAATATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTA 561

Qy        359 SerLeuLeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsn 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        562 TCTCTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAAT 621

Qy        379 AlaGlyAspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPhe 398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        622 GCTGGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTT 681

Qy        399 ValValAlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeu 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        682 GTTGTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTG 741

Qy        419 Leu 419
              |||
Db        742 CTT 744
```

RESULT 6
CO771206
LOCUS CO771206 823 bp mRNA linear EST 04-AUG-2004
DEFINITION testis_EST02303 Testis cDNA Library Gallus gallus cDNA 3', mRNA sequence.
ACCESSION CO771206
VERSION CO771206.1 GI:50972473
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 823)
AUTHORS Shin,J.
TITLE Testis cDNA Library
JOURNAL Unpublished (2004)
COMMENT Contact: JiHye Shin
Division of Animal Genetic Engineering
School of Agricultural Biotechnology, Seoul National University
San 56-1,Sillim-dong, Gwanak-gu, Seoul, 151-742, Korea
Tel: +82-31-479-7557
Fax: +82-31-479-7550
Email: paecilojh@hanmail.net
Seq primer: T3.
FEATURES Location/Qualifiers
source 1. .823
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Korean Native chicken"
/db_xref="taxon:9031"
/sex="Male"
/tissue_type="Testis"
/dev_stage="26weeks"
/clone_lib="Testis cDNA Library"
/note="Vector: Uni-ZAP XR vector; Site_1: XhoI; Site_2: EcoRI; The libraries were synthesized using the ZAP(r)-cDNA synthesis method (Stratagene). cDNA was prepared, size-fractionated and inserted into the Uni-ZAP XR vector using XhoI linker-primer and EcoRI adaptor. After in vivo excision with E. coli strain SOLR, cDNA libraries from testis contain each insert size range of 0.5~3kb (n=20), and 0.5~3.5kb (n=20)."
ORIGIN

Alignment Scores:

Pred. No.:	1.68e-111	Length:	823
Score:	1169.00	Matches:	229
Percent Similarity:	84.38%	Conservative:	14
Best Local Similarity:	79.51%	Mismatches:	23
Query Match:	52.99%	Indels:	22
DB:	7	Gaps:	2

US-10-798-532-4 (1-419) x CO771206 (1-823)

```
Qy         47 GlySerProPro---------GlyAlaLeuArgGlnPheLeuProProGlyThrThrGly 63
              |||:::||||||         ||||||         |||||||||
Db          6 GGAGCTCCACCGCGGTGGCNGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAAT 65

Qy         64 AlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeuGluGlnLeu 83
              :::|||
Db         66 TCGGCACGAGGC---------------------------------------------- 77

Qy         84 GlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSerProAsn 103
                       ||||||||||||||||||||||||||||||||||||    |||:::|||    |||
Db         78 ---------CGCAACATCGTGACTCTGTACGGCGTGTTCACCAACCACTACTCGGCCAAC 128

Qy        104 ValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeu 123
                  |||||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db        129 GGCCCGTCCCGCTGCCTGCTGCTGGAGCTGCTGGATATCAGCGTGTCTGAGCTGCTGCTG 188

Qy        124 TyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArgAspValLeu 143
              :::||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 CACTCCAGCAACCAGGGCTGCTCCATGTGGATGATCCAGCACTGCGCCCGCGACGTGCTG 248

Qy        144 GluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLysProArgAsn 163
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        249 GAAGCCCTGGCCTTCCTGCACCACAAAGGCTACGTGCACGCAGACCTGAAGCCACGCAAC 308

Qy        164 IleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLys 183
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db        309 ATCCTGTGGAGCGCCGAGGAGGAGTGCTTTAAGCTCATTGACTTTGGACTTAGCTTCAAA 368

Qy        184 GluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGlu 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GAGGGGAATCAGGATGTGAAATATATTCAAACAGACGGGTATCGGGCTCCAGAGGCAGAA 428

Qy        204 LeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaVal 223
              ||||||||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db        429 CTGCAGAACTGCCTAGCACAGGCAGGGCTCCAGAGTGAGACGGAATGTACCTCTGCTGTG 488

Qy        224 AspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLysLeuLysHis 243
              ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        489 GATCTGTGGAGTCTGGGAATTGTTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACAT 548

Qy        244 ThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAla 263
              ||||||:::|||||||||||||||   ||||||||||||||||||||   |||||||||
Db        549 ACAGTCCAATCTCAGGAATGGAAGACAAACAGTTCTGCCATCATCGATCGCATTTTTGCC 608

Qy        264 SerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSer 283
              |||:::   |||||||||:::|||||||||||||||||||||||||||||||||||||||
Db        609 AGTGAAGGGGTGGTTAATTCAGCCATTCCAGCTTATCACCTCAGAGACCTTATTAAAAGC 668

Qy        284 MetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhe 303
              |||||||||   |||      :::|||      ||||||   ||||||||||||||||||||
Db        669 ATGCTTCATTGTGACCAAGGCAAGCGAGCCTCTGCTGAAAAGGCTTTATGCAGCCCATTC 728

Qy        304 PheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrProValLeu 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 TTCAGCATTCCCTTTGCTCCCCATATTGAAGATTTGGTGATGCTCCCCACGCCTGTGCTG 788

Qy        324 ArgLeuLeuAsnValLeuAspAsp 331
              ||||||||||||||||||:::|||
Db        789 AGGCTGCTGAATGTTCTAAACGAT 812
```

RESULT 7
BU447831
LOCUS BU447831 779 bp mRNA linear EST 29-NOV-2002
DEFINITION 603767571F1 CSEQRBN13 Gallus gallus cDNA clone ChEST695n17 5', mRNA sequence.
ACCESSION BU447831
VERSION BU447831.1 GI:25937142
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 779)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .779
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Layer"
                     /db_xref="taxon:9031"
                     /clone="ChEST695n17"
                     /sex="Female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQRBN13"
                     /note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
                     EcoRI; Site_2: Not1; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with
                     EcoRI, size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer
                     reannealing hybridization was used."
ORIGIN

Alignment Scores:
Pred. No.:              8.36e-111       Length:         779
Score:                  1162.00         Matches:        224
Percent Similarity:     93.02%          Conservative:   16
Best Local Similarity:  86.82%          Mismatches:     18
Query Match:            52.67%          Indels:         1
DB:                     5               Gaps:           0

US-10-798-532-4 (1-419) x BU447831 (1-779)

Qy        158 AspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAsp 177
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db          6 GACCTGAAGCCACGCAACATCCTGTGGAGCGCGGAGGAGGAGTGCTTTAAGCTCATTGAC 65

Qy        178 PheGlyLeuSerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyr 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TTTGGACTTAGCTTCAAAGAGGGGAATCAGGATGTGAAATATATTCAAACAGACGGGTAT 125

Qy        198 ArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThr 217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        126 CGGGCTCCAGAGGCAGAACTGCAGAACTGCCTAGCACAGGCAGGGCTCCAGAGTGAGACG 185

Qy        218 GluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSer 237
              ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        186 GAATGTACCTCTGCTGTGGATCTGTGGAGTCTGGGAATTGTTTTACTGGAAATGTTCTCA 245

Qy        238 GlyMetLysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIle 257
              ||||||||||||||||||||||||:::|||||||||||||||   ||||||||||||||||

Db        246 GGAATGAAACTGAAACATACAGTCCAATCTCAGGAATGGAAGACAAACAGTTCTGCCATC 305

Qy        258 IleAspHisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeu 277
              ||||||   ||||||||||||:::   |||||||||:::||||||||||||||||||||
Db        306 ATCGATCGCATTTTTGCCAGTGAAGGGGTGGTTAATTCAGCCATTCCAGCTTATCACCTC 365

Qy        278 ArgAspLeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMet 297
              ||||||||||||||||||||||||||||  |||     :::|||     ||||||
Db        366 AGAGACCTTATTAAAAGCATGCTTCATTGTGACCAAGGCAAGCGAGCCTCTGCTGAAAAG 425

Qy        298 AlaLeuCysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMet 317
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        426 GCTTTATGCAGCCCATTCTTCAGCATTCCCTTTGCTCCCCATATTGAAGATTCGGTGATG 485

Qy        318 LeuProThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGlu 337
              ||||||||||||||||||||||||||||||||||||||||||  |||     |||   :::|||
Db        486 CTCCCCACGCCTGTGCTGAGGCTGCTGAATGTTCTAAGCGATGCTTCTCTGCAGAGTGAA 545

Qy        338 GluGluTyrGluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProVal 357
              ||||||||||||||:::||||||:::|||||||||||||||||||||||||||||||||
Db        546 GAAGAATACGAAGATATCCTGGAAGACATAAGGGAGGAGTGTCAGAAATATGGACCGGTG 605

Qy        358 ValSerLeuLeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAla 377
              ||||||||||||:::|||||||||||||||||:::|||||||||||   |||||||||
Db        606 GTTTCCTTGCTTATTCCGAAGGAAAATCCTGGTAAAGGCCAAGTCTTTGT-GAATATGCA 664

Qy        378 AsnAlaGlyAspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLys 397
              ||||||||||||||||||||||||||||:::|||||||||:::||||||||||||||||
Db        665 AATGCTGGTGATTCCAAAGCTGCCCAAAAAATGCTGACTGGAAAGATTTTTGATGGCAAG 724

Qy        398 PheValValAlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyr 415
              ||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        725 TCTGTCGTGGCTACGTNTTACCCACTGAGTGCCTATAAGAGAGGATATCTGTAC 778

RESULT 8
CO773091
LOCUS       CO773091                 896 bp    mRNA    linear   EST 04-AUG-2004
DEFINITION  testis_EST04188 Testis cDNA Library Gallus gallus cDNA 3', mRNA
            sequence.
ACCESSION   CO773091
VERSION     CO773091.1  GI:50974358
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 896)
  AUTHORS   Shin,J.
  TITLE     Testis cDNA Library
  JOURNAL   Unpublished (2004)
COMMENT     Contact: JiHye Shin
            Division of Animal Genetic Engineering
            School of Agricultural Biotechnology, Seoul National University
            San 56-1,Sillim-dong, Gwanak-gu, Seoul, 151-742, Korea
            Tel: +82-31-479-7557
            Fax: +82-31-479-7550
            Email: paecilojh@hanmail.net
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .896
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Korean Native chicken"
                     /db_xref="taxon:9031"
                     /sex="Male"
                     /tissue_type="Testis"
                     /dev_stage="26weeks"
                     /clone_lib="Testis cDNA Library"
                     /note="Vector: Uni-ZAP XR vector; Site_1: XhoI; Site_2:
                     EcoRI; The libraries were synthesized using the
                     ZAP(r)-cDNA synthesis method (Stratagene). cDNA was
                     prepared, size-fractionated and inserted into the Uni-ZAP
```

XR vector using XhoI linker-primer and EcoRI adaptor. After in vivo excision with E. coli strain SOLR, cDNA libraries from testis contain each insert size range of 0.5~3kb (n=20), and 0.5~3.5kb (n=20)."

ORIGIN

Alignment Scores:

Pred. No.:	2.71e-110	Length:	896
Score:	1158.00	Matches:	240
Percent Similarity:	81.59%	Conservative:	17
Best Local Similarity:	76.19%	Mismatches:	31
Query Match:	52.49%	Indels:	27
DB:	7	Gaps:	2

US-10-798-532-4 (1-419) x CO773091 (1-896)

```
Qy         46 ProGlySer-ProPro---------GlyAlaLeuArgGlnPheLeuProProGlyThrTh 62
              ||||||::: ||||||          ||||||          |||||||||
Db         11 CCTGGNACCTCCACCGCGGTGGCGNGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAG 70

Qy         62 rGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeuGluGl 82
                  :::|||
Db         71 GAATTCGGCACGAGGC---------------------------------------------- 86

Qy         82 nLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSerPr 102
                        ||||||||||||||||||||||||||||||||||   |||:::|||
Db         87 ------------CGCAACATCGTGACTCTGTACGGCGTGTTCACCAACCACTACTCGGC 133

Qy        102 oAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGluLeuLe 122
               |||   |||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        134 CAACGGCCCGTCCCGCTGCCTGCTGCTGGAGCTGCTGGATATCAGCGTGTCTGAGCTGCT 193

Qy        122 uLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArgAspVa 142
              ||||:::||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCTGCACTCCAGCAACCAGGGCTGCTCCATGTGGATGATCCAGCACTGCGCCCGCGACGT 253

Qy        142 lLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLysProAr 162
              ||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db        254 GCTGGAAGCCCTGGCCTTCCTGCACCACAAAGGCTACGTGCACGCAGACCTGAAGCCACG 313

Qy        162 gAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeuSerPh 182
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        314 CAACATCCTGTGGAGCGCCGAGGAGGAGTGCTTTAAGCTCATTGACTTTGGACTTAGCTT 373

Qy        182 eLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaProGluAl 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 CAAAGAGGGGAATCAGGATGTGAAATATATTCAAACAGACGGGTATCGGGCTCCAGAGGC 433

Qy        202 aGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAl 222
              ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        434 AGAACTGCAGAACTGCCTAGCACAGGCAGGGCTCCAGAGTGAGACGGAATGTACCTCTGC 493

Qy        222 aValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLysLeuLy 242
              |||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        494 TGTGGATCTGTGGAGTCTGGGAATTGTTTTACTGGAAATGTTCTCAGGAATGAAACTGAA 553

Qy        242 sHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePh 262
              ||||||||||:::|||||||||||||||   ||||||||||||||||||||   |||||
Db        554 ACATACAGTCCAATCTCAGGAATGGAAGACAAACAGTTCTGCCATCATCGATCGCATTTT 613

Qy        262 eAlaSer-LysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIleL 282
              ||||||| |||   |||||||||:::|||||||||||||||||||||||||||||||||||
Db        614 TGCCAGTGAAGGGGGTGGTTAATTCAGCCATTCCAGCTTATCACCTCAGAGACCTTATTA 673

Qy        282 ysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerP 302
              |||||||||||||||   |||      :::|||      ||||||   ||||||||||||||
Db        674 AAAGCATGCTTCATTGTGACCAAGGCAAGCGAGCCTCTGCTGAAAAGGCTTTATGCAGCC 733

Qy        302 roPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrProV 322
              ||||||||||||||||||||||||||||||:::::||||||||||||||||||||||||||
Db        734 CATTCTTCAGCATTCCCTTTGCTCCCCATATGAAAGATTTGGTGATGCTCCCCACGCCTG 793

Qy        322 alLeuArgLeu-LeuAsnValLeuAspAspAspTyrLeuGlyAsn-GluGluGluTyrGl 341
```

```
                ||||||||||| |||||||||||   |||      |||   ::: ||||||||||||||
Db        794 TGCTGAGGCTGNCTGAATGTTCTAAGCGATGCTTCTCTGCAGAGTGGAAGAAGAATACGA 853

Qy        341 uAspValValGlu-AspValLysGluGluCysGln 352
                  |:::   :::|||          ||||||||||||
Db        854 AGAATCCCTGGAGACATAGGGNAGAGGAGTGTCAG 888
```

RESULT 9
CN232453

LOCUS	CN232453 781 bp mRNA linear EST 09-APR-2004
DEFINITION	WLB074D11.ab1 WLtestis Gallus gallus cDNA 5', mRNA sequence.
ACCESSION	CN232453
VERSION	CN232453.1 GI:46336197
KEYWORDS	EST.
SOURCE	Gallus gallus (chicken)
ORGANISM	Gallus gallus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE	1 (bases 1 to 781)
AUTHORS	Savolainen,P., Fitzsimmons,C.J., Arvestad,L., Andersson,L. and Lundeberg,J.
TITLE	EST analysis of brain and testis cDNA libraries from White leghorn and Red Jungle Fowl
JOURNAL	Unpublished (2004)
COMMENT	Contact: Peter Savolainen Department of Biotechnology Royal Institute of Technology, KTH SE-106 91 Stockholm, SWEDEN Tel: +46 (0)8 5537 8481 Fax: +46 (0)8 5537 8335 Email: Peter.Savolainen@biotech.kth.se Seq primer: M13 reverse primer.

FEATURES	Location/Qualifiers
source	1. .781
	/organism="Gallus gallus"
	/mol_type="mRNA"
	/strain="White Leghorn"
	/db_xref="taxon:9031"
	/sex="male"
	/lab_host="ElectroMAX DH10B (Invitrogen)"
	/clone_lib="WLtestis"
	/note="Organ: testis; Vector: pSPORT-1; Site_1: Hind III; Site_2: EcoRI; The cDNA libraries were created with the Superscript Plasmid System (Invitrogen)."

ORIGIN

Alignment Scores:

Pred. No.:	9.14e-108	Length:	781
Score:	1133.00	Matches:	225
Percent Similarity:	93.10%	Conservative:	18
Best Local Similarity:	86.21%	Mismatches:	18
Query Match:	51.36%	Indels:	3
DB:	7	Gaps:	0

US-10-798-532-4 (1-419) x CN232453 (1-781)

```
Qy        109 LeuLeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGln 128
              ||||||||||||||||||||||:::|||||||||||||||||||||:::|||||||:::|||
Db          2 CTGCTGCTGGAGCTGCTGGATATCAGCGTGTCTGAGCTGCTGCTGCACTCCAGCAACCAG 61

Qy        129 GlyCysSerMetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPhe 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GGCTGCTCCATGTGGATGATCCAGCACTGCGCCCGCGACGTCCTGGAAGCCCTGGCCTTC 121

Qy        149 LeuHisHisGluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAla 168
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CTGCACCACAAAGGCTACGTGCACGCAGACCTGAAGCCACGCAACATCCTGTGGAGCGCG 181

Qy        169 GluAsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAsp 188
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GAGGAGGAGTGCTTTAAGCTCATTGACTTTGGACTTAGCTTCAAAGAGGGGAATCAGGAT 241
```

```
Qy        189 ValLysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeu 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GTGAAATATATTCAAACAGACGGGTATCGGGCTCCAGAGGCAGAACTGCAGAACTGCCTA 301

Qy        209 AlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeu 228
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        302 GCACAGGCAGGGCTCCAGAGTGAGACGGAATGTACCTCTGCTGTGGATCTGTGGAGTCTG 361

Qy        229 GlyIleIleLeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGln 248
              ||||||:::||||||||||||||||||||||||||||||||||||:::||||||||||||||
Db        362 GGAATTGTTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCCAATCTCAG 421

Qy        249 GluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAlaValVal 268
              |||||||||   ||||||||||||||||||||   ||||||||||:::   ||||||
Db        422 GAATGGAAGACAAACAGTTCTGCCATCATCGATCGCATTTTTGCCAGTGAAGGGGTGGTT 481

Qy        269 AsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAsp 288
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        482 AATTCAGCCATTCCAGCTTATCACCTCAGAGACCTTATTAAAAGCATGCTTCATTGTGAC 541

Qy        289 ProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPhe 308
                    :::|||      ||||||   ||||||||||||||||||||||||||||||||
Db        542 CAAGGCAAACGAGCCTCTGCTGAAAAGGCTTTATGCAGCCCATTCTTCAGCATTCCCTTT 601

Qy        309 AlaProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeuAsnVal 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GCTCCCCATATTGAAGATTTGGTGATGCTCCCCACGCCTGTGCTGAGGCTGCTGAATGTT 661

Qy        329 LeuAspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAspValLys 348
              |||   |||      |||   :::|||   |||||||||||:::::|||   ::::::
Db        662 CTAAGCGATGCTTCTCTGCAGAGTGAAGA-GAATACGAAGATATCCTGGAAGA-ATAAGG 719

Qy        349 GluGluCysGlnLysTyrGlyProValValSerLeuLeuValProLysGluAsnProGly 368
              ||||||||||||||||||||||||   |||||||||:::|||:::   |||||||||
Db        720 GAGGAGTGTCAGAAATATGGACCGGTGGGTTCCTTGCTTATTCCGGAGGA-AATCCTGGT 778

Qy        369 Arg 369
              :::
Db        779 AAA 781

RESULT 10
BU541306
LOCUS       BU541306                 932 bp    mRNA    linear   EST 13-SEP-2002
DEFINITION  AGENCOURT_10327419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6572970
            5', mRNA sequence.
ACCESSION   BU541306
VERSION     BU541306.1  GI:22851747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 932)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2765  row: o  column: 18
            High quality sequence stop: 631.
FEATURES             Location/Qualifiers
     source          1. .932
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6572970"
                     /tissue_type="carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_40"
                     /note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed  by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California,  Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life  Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                6.3e-105        Length:          932
Score:                    1107.00         Matches:         227
Percent Similarity:       87.22%          Conservative:    5
Best Local Similarity:    85.34%          Mismatches:      18
Query Match:              50.18%          Indels:          16
DB:                       5               Gaps:            2

US-10-798-532-4 (1-419) x BU541306 (1-932)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 208

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 268

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        269 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 328

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 388

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 448

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 508

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        509 TTGCTCTTATATTCCAGTCACCAGCGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 568

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeu-Ly 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        569 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCCAA 628

Qy        160 sProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLe 180
              ||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        629 ACCACGTAACATATTGTGGAGTGCAAAGAATGAATGTTTTAAACTCATTGACTTTGGACT 688

Qy        180 u-SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaP 200
              | ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        689 TTAGCTTCAAAGAAAGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTC 748

Qy        200 roGlu-AlaGluLeuGlnAsnCysLeuAlaGlnAla-GlyLeuGlnSerAspThrGluCy 219
              |||||    ||||||||||||||||||||   ||| ||||||||||||||||:::||
Db        749 CAGAAAACAGAATTGCAAAATTGCTTGGCCCANGCTTGGCCTGCAAAGTGATACAAAATG 808

Qy        219 sThrSer-AlaValAsp-LeuTrpSer------LeuGlyIleIleLeuLeuGluMetPhe 236
              ||||||| ||||||||| |||||||||      |||
Db        809 TACCTCCAGCTGTTGATCTTGTGGAGCCCTAAGAAATCATTTTACCTGGAAAATGGTCTC 868

Qy        237 SerGlyMetLysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAla 256
                ||||||                          :::   ||||||         |||
```

```
Db        869 CAGGGAATG----------------------GAAAACTGGAAAACCTACCGGTCCAGA 904

Qy        257 IleIle 258
              |||:::
Db        905 ATCCTC 910

RESULT 11
BU385761
LOCUS       BU385761                717 bp    mRNA    linear   EST 28-NOV-2002
DEFINITION  603582650F1 CSEQCHN75 Gallus gallus cDNA clone ChEST533d13 5', mRNA
            sequence.
ACCESSION   BU385761
VERSION     BU385761.1  GI:25893762
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 717)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .717
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="White Leghorn, Hisex"
                     /db_xref="taxon:9031"
                     /clone="ChEST533d13"
                     /dev_stage="36"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN75"
                     /note="Organ: trunks; Vector: pBluescript II KS(+);
                     Site_1: EcoRI; Site_2: Not1; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with
                     EcoRI, size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer
                     reannealing hybridization was used."
ORIGIN

Alignment Scores:
Pred. No.:              5.41e-105      Length:          717
Score:                  1106.00        Matches:         210
Percent Similarity:     93.72%         Conservative:    14
Best Local Similarity:  87.87%         Mismatches:      14
Query Match:            50.14%         Indels:          1
DB:                     5              Gaps:            0

US-10-798-532-4 (1-419) x BU385761 (1-717)

Qy        134 MetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeuHisHisGluGly 153
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db          3 ATGATCCAGCACTGCGCCCGCGACGTCCTGGAAGCCCTGGCCTTCCTGCACCACAAAGGC 62

Qy        154 TyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGluCysPhe 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db         63 TACGTGCACGCAGACCTGAAGCCACGCAACATCCTGTGGAGCGCGGAGGAGGAGTGCTTT 122

Qy        174 LysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspValLysTyrIleGln 193
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 AAGCTCATTGACTTTGGACTTAGCTTCAAAGAGGGGAATCAGGATGTGAAATATATTCAA 182

Qy        194 ThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeu 213
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ACAGACGGGTATCGGGCTCCAGAGGCAGAACTGCAGAACTGCCTAGCACAGGCAGGGCTC 242

Qy        214 GlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeu 233
              ||||||:::|||||||||||||||||||||||||||||||||||||||:::||||||
Db        243 CAGAGTGAGACGGAATGTACCTCTGCTGTGGATCTGTGGAGTCTGGGAATTGTTTTACTG 302

Qy        234 GluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsn 253
              |||||||||||||||||||||||||||||||:::|||||||||||||||   |||
Db        303 GAAATGTTCTCAGGAATGAAACTGAAACATACAGTCCAATCTCAGGAATGGAAGACAAAC 362

Qy        254 SerSerAlaIleIleAspHisIlePheAlaSerLysAlaValValAsnAlaAlaIlePro 273
              ||||||||||||||||||   |||||||||||:::   |||||||||:::|||||||||
Db        363 AGTTCTGCCATCATCGATCGCATTTTTGCCAGTGAAGGGGTGGTTAATTCAGCCATTCCA 422

Qy        274 AlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspProSerArgArgIle 293
              |||||||||||||||||||||||||||||||||||||   |||      :::|||
Db        423 GCTTATCACCTCAGAGACCTTATTAAAAGCATGCTTCATTGTGACCAAGGCAAACGAGCC 482

Qy        294 ProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaProHisIleGlu 313
                 ||||||   |||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TCTGCTGAAAAGGCTTTATGCAGCCCATTCTTCAGCATTCCCTTTGCTCCCCATATTGAA 542

Qy        314 AspLeuValMetLeuProThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyr 333
              |||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        543 GATTTGGTGATGCTCCCCACGCCTGTGCTGAGGCTGCTGAATGTTCTAAGCGATGCTTCT 602

Qy        334 LeuGlyAsnGluGluGluTyrGluAspValValGluAspValLysGluGluCysGlnLys 353
              |||   :::|||||||||||||||||||:::::|||||:::::|||||||||||||||||
Db        603 CTGCAGAGTGAAGAAGAATACGAAGATATCCTGGAAGACATAAGGGAGGAGTGTCAGAAA 662

Qy        354 TyrGlyProValValSerLeuLeuValProLysGlu-AsnProGlyArgGlyGln 371
              |||||||||||||||||||||||:::|||||||||| |||||||||:::||||||
Db        663 TATGGACCGGTGGTTTCCTTGCTTATTCCGAAGGAACAATCCTGGTAAAGGCCAA 717

RESULT 12
BE895119
LOCUS       BE895119                899 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601436041F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920939 5',
            mRNA sequence.
ACCESSION   BE895119
VERSION     BE895119.1  GI:10358193
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 899)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9753  row: b  column: 12
            High quality sequence stop: 626.
FEATURES             Location/Qualifiers
```

```
     source          1. .899
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3920939"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

Alignment Scores:
Pred. No.:              2.83e-103       Length:         899
Score:                  1091.00         Matches:        210
Percent Similarity:     99.53%          Conservative:   0
Best Local Similarity:  99.53%          Mismatches:     0
Query Match:            49.46%          Indels:         1
DB:                     2               Gaps:           0

US-10-798-532-4 (1-419) x BE895119 (1-899)

Qy        113 LeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGlnGlyCysSerMet 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATG 60

Qy        133 TrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeuHisHisGlu 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAG 120

Qy        153 GlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGluCys 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGT 180

Qy        173 PheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspValLysTyrIle 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATT 240

Qy        193 GlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGly 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGC 300

Qy        213 LeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeu 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTA 360

Qy        233 LeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGluTrpLysAla 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCA 420

Qy        253 AsnSerSerAlaIleIleAspHisIlePheAlaSerLysAlaValValAsnAlaAlaIle 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATT 480

Qy        273 ProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspProSerArgArg 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGA 540

Qy        293 IleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaProHisIle 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATT 600

Qy        313 GluAspLeuValMetLeu-ProThrProVal 322
              |||||||||||||||||| ||||||||||||
Db        601 GAAGATCTGGTCATGCTTTCCCACTCCAGTT 631

RESULT 13
BI562420
LOCUS       BI562420                 772 bp    mRNA    linear   EST 05-SEP-2001
DEFINITION  603256316F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
            mRNA sequence.
ACCESSION   BI562420
VERSION     BI562420.1  GI:15449747
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 772)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11755  row: p  column: 03
            High quality sequence stop: 727.
FEATURES             Location/Qualifiers
     source          1. .772
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5298674"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTVN-3', size-selected for  average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation).  Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              4.06e-102       Length:         772
Score:                  1079.00         Matches:        212
Percent Similarity:     96.82%          Conservative:   1
Best Local Similarity:  96.36%          Mismatches:     5
Query Match:            48.91%          Indels:         3
DB:                     4               Gaps:           1

US-10-798-532-4 (1-419) x BI562420 (1-772)

Qy         92 LeuTyrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeuLeuLeu 111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 TTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGTCTGTTGCTT 167

Qy        112 GluLeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGlnGlyCysSer 131
              |||||||||||||||||||||||||   |||||||||||||||||   |||||||||
Db        168 GAACTCCTGGATGTCAGTGTTTCGGAANTGCTCTTATATTCCAGTCACCANGGTTGTTCC 227

Qy        132 MetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeuHisHis 151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 ATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCAT 287

Qy        152 GluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGlu 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCAGAGAATGAA 347

Qy        172 CysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspValLysTyr 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 TGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 407
```

```
Qy        192 IleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAla 211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 ATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCT 467

Qy        212 GlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIle 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        468 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 527

Qy        232 LeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGluTrpLys 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 587

Qy        252 AlaAsnSerSerAlaIleIleAspHisIle-PheAlaSerLysAlaValValAsnAlaAl 271
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        588 GCAAACAGTTCTGCTATTATTGATCACATTATTTGCCAGTAAAGCAGTGGTGAATGCCGC 647

Qy        271 aIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspProSerAr 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 AATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAG 707

Qy        291 gArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaPro 310
              |||||||||||||||||||||||||||||  ||||||   :::      |||||||||
Db        708 AAGAATTCCTGCTGAAATGGCATTGTGCAG-CCATTCTTAGCA---TCCTTTGCCCCT 761

RESULT 14
BU205357
LOCUS       BU205357                794 bp    mRNA    linear   EST 25-NOV-2002
DEFINITION  603104569F1 CSEQCHN03 Gallus gallus cDNA clone ChEST41j18 5', mRNA
            sequence.
ACCESSION   BU205357
VERSION     BU205357.1  GI:25370848
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 794)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .794
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="White Leghorn, Hisex"
                     /db_xref="taxon:9031"
                     /clone="ChEST41j18"
                     /tissue_type="whole embryo"
                     /dev_stage="20-21"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN03"
                     /note="Organ: whole embryo; Vector: pBluescript II KS(+);
                     Site_1: EcoRI; Site_2: Not1; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with
                     EcoRI, size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer
                     reannealing hybridization was used."
ORIGIN

Alignment Scores:
Pred. No.:                1.11e-101      Length:          794
Score:                    1075.00        Matches:         221
Percent Similarity:       89.10%         Conservative:    16
Best Local Similarity:    83.08%         Mismatches:      27
Query Match:              48.73%         Indels:          4
DB:                       5              Gaps:            0

US-10-798-532-4 (1-419) x BU205357 (1-794)

Qy         82 GlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSer 101
              ||||||:::|||||||||||||||||||||||||||||||||||||   |||:::|||
Db          1 CAGCTCCGCGGGCACCGCAACATCGTGACTCTGTACGGCGTGTTCACCAACCACTACTCG 60

Qy        102 ProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGluLeu 121
                 |||  ||||||||||||||||||||||||||||||:::|||||||||||||||
Db         61 GCCAACGGCCCGTCCCGCTGCCTGCTGCTGGAGCTGCTGGATATCAGCGTGTCTGAGCTG 120

Qy        122 LeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArgAsp 141
              ||||||:::|||||||:::|||  |||||||||||||||||||||||||||||||||||||
Db        121 CTGCTGCACTCCAGCAACCAGGTCTGCTCCATGTGGATGATCCAGCACTGCGCCCGCGAC 180

Qy        142 ValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLysPro 161
              ||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db        181 GTCCTGGAAGCCCTGGCCTTCCTGCACCACAAAGGCTACGTGCACGCAGACCTGAAGCCA 240

Qy        162 ArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeuSer 181
              |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        241 CGCAACATCCTGTGGAGCGCGGAGGAGGAGTGCTTTAAGCTCATTGACTTTGGACTTAGC 300

Qy        182 Phe-LysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaProGl 201
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCAAAAGAGGGGAATCAGGATGTGAAATATATTCAAACAGACGGGTATCGGGCTCCAGA 360

Qy        201 uAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSe 221
              ||||||||||||||||||||||||||||||||||||||||||::||||||||||||||||
Db        361 GGCAGAACTGCAGAACTGCCTAGCACAGGCAGGGCTCCAGAGTGAGACGGAATGTACCTC 420

Qy        221 rAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLysLe 241
              ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db        421 TGCTGTGGATCTGTGGAGTCTGGGAATTGTTTTACTGGAAATGTTCTCAGGAATGAAACT 480

Qy        241 uLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIl 261
              ||||||||||||:::|||||||||||||||  ||||||||||||||||||||||||   ||
Db        481 GAAACATACAGTCCAATCTCAGGAATGGAAGACAAACAGTTCTGCCATCATCGATCGCAT 540

Qy        261 ePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIl 281
              ||||||||||:::  |||||||||  ||||||||||||||||||||||||||||||||||||
Db        541 TTTGCCAGTGAAGGGGTGGTTAATCCAGCCATTCCAGCTTATCACCTCAGAGACCTTAT 600

Qy        281 eLysSerMetLeuHisAspAsp-ProSerArgArgIleProAlaGluMetAlaLeuCysS 301
              ||||||||||||||||  |||    ::::::|||      ||||||   |||   ||||
Db        601 TAAAAGCATGCTTCATTGTGACCAAGGCAAAACGAGCCTCTGCTGAAAAGGCTTAATGCA 660

Qy        301 erProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrP 321
              ||||||||||||||||||  |||  ||||||||||||||||||||||||||||||   |
Db        661 GCCCATTCTTCAGCATTCCCTTGGCTCC-CATATTGAAGATTTGGTGATGCTCCCAAG-C 718

Qy        321 roValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyrG 341
              ||||||||||||||||||||||||||||||          :::                :::||||
Db        719 CTGTGCTGAGGCTGCTGAATGTTCTAAGGCGAAGGTTCTCTGCAGAGTGAAGAAAATACG 778

Qy        341 luAspValValGlu 345
              ||::::::   |||
Db        779 AGAATATCTGGGAA 792

RESULT 15
```

```
CF593363
LOCUS       CF593363                834 bp    mRNA    linear   EST 26-SEP-2003
DEFINITION  AGENCOURT_15622279 NIH_MGC_147 Homo sapiens cDNA clone
            IMAGE:30531268 5', mRNA sequence.
ACCESSION   CF593363
VERSION     CF593363.1  GI:36346826
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 834)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Stefan Hansson
             cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
            and advice from Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM621  row: m  column: 05
            High quality sequence stop: 660.
FEATURES             Location/Qualifiers
     source          1. .834
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30531268"
                     /tissue_type="Human Placenta"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_147"
                     /note="Organ: placenta; Vector: pBluescriptR; Site_1:
                     alI-XhoI; Site_2: BamH; Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTTVN-3', size-selected  for average
                     insert size 2.3 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using  the Cap-trapper method (Carninci, in
                     preparation). Library constructed  by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: This is
                     a NIH_MGC library."
ORIGIN
```

Alignment Scores:

Pred. No.:	3.54e-101	Length:	834
Score:	1070.50	Matches:	213
Percent Similarity:	95.13%	Conservative:	2
Best Local Similarity:	94.25%	Mismatches:	5
Query Match:	48.53%	Indels:	6
DB:	7	Gaps:	1

US-10-798-532-4 (1-419) x CF593363 (1-834)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 214

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 274

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        275 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 334

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 394

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 454

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 514

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        515 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 574

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        575 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 634

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGly-Le 180
              ||||||||||||||||||||||||||||||||||   |||||||||||||| ||
Db        635 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTANACTCATTGACTTTGGACTT 694

Qy        180 uSerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArg-AlaP 200
              |   |||||||||   ||||||||||||||||||||||||||||||||||||| ||||
Db        695 AGGCTTCAAAGAAAGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGGCTC 754

Qy        200 roGluAlaGluLeuGln-AsnCysLeuAla---GlnAlaGly-LeuGlnSerAspThrGl 218
              |||||||||||||||| |||||||||      :::   ||| ||||||||||||||||
Db        755 CAGAAGCAGAATTGCAAAAATTGCTTGGGCCCCAGGCTTGGCCCTGCAGAGTGATACAGA 814

Qy        218 u-CysThr 220
              | ||||||
Db        815 AATGTACC 822
```

Search completed: November 29, 2004, 13:27:52
Job time : 3168.5 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 502 Seconds
(without alignments)
4381.489 Million cell updates/sec

Title: US-10-798-532-4
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAFG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10798532/runat_29112004_112452_15882/app_query.fasta_1.1166
-DB=N_Geneseq_23Sep04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532_@CGN_1_1_886_@runat_29112004_112452_15882 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2206	100.0	1260	3	AAZ51356	Aaz51356 Human tra
2	2203	99.9	1260	3	AAZ51355	Aaz51355 Human wil
3	2195	99.5	1260	4	AAI66829	Aai66829 Human pro
4	2195	99.5	2008	6	ABQ99383	Abq99383 Human cod
5	2195	99.5	2622	4	AAI66828	Aai66828 Human pro
6	2195	99.5	4065	4	AAH72714	Aah72714 Human cer
7	2195	99.5	4065	5	ADL45238	Adl45238 Human ova
8	2188	99.2	1260	4	AAS06734	Aas06734 Polynucle
9	2188	99.2	1824	4	AAD18818	Aad18818 Human kin
10	2183	99.0	3244	10	ADB52857	Adb52857 Primary r
11	2042.5	92.6	3998	12	ADQ24401	Adq24401 Human sof
12	789	35.8	529	4	AAH70764	Aah70764 Human cer
13	766.5	34.7	722	3	AAA02535	Aaa02535 Human col
14	729	33.0	470	4	AAH71966	Aah71966 Human cer
15	636	28.8	461	4	AAH69101	Aah69101 Human cer
16	559.5	25.4	411	6	ABQ99225	Abq99225 Human ORF
17	361	16.4	231	3	AAA01259	Aaa01259 Human col
18	311.5	14.1	419	5	ADI73740	Adi73740 Human ova
19	311.5	14.1	419	5	ADI67345	Adi67345 Human ova
c 20	284	12.9	546	6	ABQ53127	Abq53127 Oligonucl
21	284	12.9	546	6	ABQ53126	Abq53126 Oligonucl
c 22	284	12.9	547	6	ABQ44465	Abq44465 Oligonucl
23	284	12.9	547	6	ABQ44464	Abq44464 Oligonucl
24	255	11.6	546	6	ABQ53129	Abq53129 Oligonucl
c 25	255	11.6	546	6	ABQ53128	Abq53128 Oligonucl
26	255	11.6	547	6	ABQ44463	Abq44463 Oligonucl
c 27	255	11.6	547	6	ABQ44462	Abq44462 Oligonucl
28	250	11.3	442	9	ACH28557	Ach28557 Human adu
29	249	11.3	1065	10	ADD21399	Add21399 Golden ha
30	246.5	11.2	1400	10	ADD21407	Add21407 Rat cycli
31	245	11.1	1708	10	ADD21395	Add21395 Mouse cyc
32	240.5	10.9	2472	3	AAC45716	Aac45716 Arabidops
33	240	10.9	1297	10	ADD21393	Add21393 Human cyc
34	240	10.9	2213	6	ABK84073	Abk84073 Human cDN
35	240	10.9	2213	10	ADD21403	Add21403 Chinese h
36	240	10.9	2213	11	ADI31711	Adi31711 Human cDN
37	240	10.9	2213	12	ADH17138	Adh17138 Human cdc
38	240	10.9	2328	10	ACC72837	Acc72837 Human can
39	240	10.9	2734	12	ADQ25034	Adq25034 Human sof
40	239	10.8	1650	12	ADI45326	Adi45326 Rice isop
41	237	10.7	1209	10	ADD21401	Add21401 Golden ha
42	236.5	10.7	1240	6	ABL88413	Abl88413 Pain regu
43	236.5	10.7	1635	2	AAV71073	Aav71073 CDK2-gree
44	236.5	10.7	1665	3	AAA90869	Aaa90869 Rat stres
45	236	10.7	1297	6	ABQ60787	Abq60787 Human 10C

ALIGNMENTS

RESULT 1
AAZ51356
ID AAZ51356 standard; DNA; 1260 BP.
XX
AC AAZ51356;
XX
DT 06-JUN-2000 (first entry)
XX
DE Human transdominant mutant serine/threonine kinase KIS (hKIS) gene.
XX
KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
KW transdominant gene; mutant; cyclin-dependent kinase inhibitor; CKI; p27;
KW modulator; treatment; cell proliferative disease; vascular disorder;
KW gene therapy; atherosclerosis; restenosis; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1260
FT /*tag= a
FT /product= "Human transdominant mutant KIS protein"
FT /transl_except= (pos:556. .561, Asn)
FT misc_feature 161
FT /*tag= b
FT /note= "A to G substitution resulting in a Lys to Arg
FT change in the amino acid sequence"
XX
PN WO200011165-A1.
XX
PD 02-MAR-2000.

```
XX
PF   20-AUG-1999;   99WO-US018903.
XX
PR   21-AUG-1998;   98US-0097710P.
XX
PA   (NABE/) NABEL G J.
PA   (NABE/) NABEL E G.
XX
PI   Nabel GJ,  Nabel EG;
XX
DR   WPI; 2000-237648/20.
DR   P-PSDB; AAY70306.
XX
PT   Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT   for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT   proliferation.
XX
PS   Claim 7; Page 59; 70pp; English.
XX
CC   The present DNA sequence encodes a transdominant mutant human KIS (hKIS),
CC   constructed by site directed mutagenesis. A single nucleotide
CC   substitution (A to G) results in a lysine to arginine change in the
CC   protein sequence. hKIS is a serine/threonine kinase, that acts as an
CC   inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
CC   controls cell proliferation and is localised predominantly in the
CC   nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC   G1 phase. The hKIS sequences are used to modulate cell proliferation and
CC   treat cell proliferative and vascular diseases. The polynucleotide
CC   sequence may be used in gene therapy to treat vascular disorders such as
CC   restenosis or atherosclerosis
XX
SQ   Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.44e-207      Length:          1260
Score:                    2206.00        Matches:         419
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       3              Gaps:            0

US-10-798-532-4 (1-419) x AAZ51356 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAGGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257

RESULT 2
AAZ51355
ID   AAZ51355 standard; DNA; 1260 BP.
XX
AC   AAZ51355;
XX
DT   06-JUN-2000  (first entry)
XX
DE   Human wild type serine/threonine kinase KIS (hKIS) gene.
XX
KW   KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
KW   dominant gene; cyclin-dependent kinase inhibitor; CKI; p27; modulator;
KW   treatment; cell proliferative disease; vascular disorder; gene therapy;
KW   atherosclerosis; restenosis; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1260
```

```
FT                   /*tag=  a
FT                   /product= "Human wild type KIS protein"
FT                   /transl_except= (pos:556. .561, Asn)
XX
PN   WO200011165-A1.
XX
PD   02-MAR-2000.
XX
PF   20-AUG-1999;   99WO-US018903.
XX
PR   21-AUG-1998;   98US-0097710P.
XX
PA   (NABE/) NABEL G J.
PA   (NABE/) NABEL E G.
XX
PI   Nabel GJ,  Nabel EG;
XX
DR   WPI; 2000-237648/20.
DR   P-PSDB; AAY70305.
XX
PT   Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT   for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT   proliferation.
XX
PS   Claim 1; Page 56; 70pp; English.
XX
CC   The present DNA sequence encodes a wild type human KIS (hKIS), isolated
CC   from a human B-cell library, using a yeast two hybrid screening system.
CC   hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
CC   cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
CC   that controls cell proliferation and is localised predominantly in the
CC   nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC   G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
CC   The hKIS sequences are used to modulate cell proliferation and treat cell
CC   proliferative and vascular diseases. The polynucleotide sequence may be
CC   used in gene therapy to treat vascular disorders such as restenosis or
CC   atherosclerosis
XX
SQ   Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.84e-207      Length:          1260
Score:                  2203.00        Matches:         418
Percent Similarity:     100.00%        Conservative:    1
Best Local Similarity:  99.76%         Mismatches:      0
Query Match:            99.86%         Indels:          0
DB:                     3              Gaps:            0

US-10-798-532-4 (1-419) x AAZ51355 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGGGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257

RESULT 3
AAI66829
ID   AAI66829 standard; cDNA; 1260 BP.
XX
AC   AAI66829;
XX
DT   07-JAN-2002  (first entry)
XX
DE   Human protein kinase polypeptide 13887 coding sequence.
XX
KW   Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KW   cytostatic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;
```

```
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1260
FT                   /*tag= a
FT                   /product= "protein kinase 13887"
XX
PN   WO200173050-A2.
XX
PD   04-OCT-2001.
XX
PF   23-MAR-2001; 2001WO-US009483.
XX
PR   24-MAR-2000; 2000US-0191846P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Meyers R;
XX
DR   WPI; 2001-611632/70.
DR   P-PSDB; AAG65767.
XX
PT   New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT   useful in diagnosis of cancer or cellular proliferation or
PT   differentiation disorders and to screen for polypeptide modulators useful
PT   to treat such conditions.
XX
PS   Claim 1; Fig 10A-C; 169pp; English.
XX
CC   The invention provides novel human protein kinase polypeptides, 3714,
CC   16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC   protein kinase polypeptides can be expressed by standard recombinant
CC   methodology.3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC   are useful for diagnostic and screening methods to identify subjects (at
CC   risk of) having cancer or cellular proliferation and/or differentiation
CC   disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC   modulators are useful for the treatment of cancer, particularly colon
CC   cancer or cellular proliferation and/or differentiation disorders. Other
CC   disorders associated with 3714, 16742, 23546 or 13887 expression or
CC   activity that can be treated include bone related disorders, inflammatory
CC   disorders, autoimmune diseases, cardiovascular disorders and liver
CC   diseases. The present sequence represents a human protein kinase
CC   polypeptide 13887 coding sequence
XX
SQ   Sequence 1260 BP; 312 A; 279 C; 334 G; 335 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.75e-206      Length:          1260
Score:                    2195.00        Matches:         417
Percent Similarity:       99.76%         Conservative:    1
Best Local Similarity:    99.52%         Mismatches:      1
Query Match:              99.50%         Indels:          0
DB:                       4              Gaps:            0

US-10-798-532-4 (1-419) x AAI66829 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257

RESULT 4
ABQ99383
ID   ABQ99383 standard; cDNA; 2008 BP.
XX
```

AC ABQ99383;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human coding sequence SEQ ID 116.
XX
KW Human; expressed sequence tag; EST; haematopoietic disorder;
KW central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200259260-A2.
XX
PD 01-AUG-2002.
XX
PF 16-NOV-2001; 2001WO-US042950.
XX
PR 17-NOV-2000; 2000US-00714936.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-590824/63.
DR N-PSDB; ABP64797.
XX
PT New isolated polynucleotide, useful in research, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
PT aberrant protein expression or biological activity.
XX
PS Claim 1; SEQ ID NO 116; 394pp; English.
XX
CC The present invention relates to novel human coding sequences (ABQ99268-
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC therapeutic, diagnostic and research methods. The polynucleotides may be
CC used in the field of molecular biology as hybridisation probes, primers
CC for PCR, for chromosome and gene mapping, for the recombinant production
CC of protein, or in generation of anti-sense DNA or RNA. The
CC polynucleotides are useful in diagnostics as expressed sequence tags
CC (ESTs) for identifying expressed genes or for physical mapping of the
CC human genome. The proteins may be used as molecular weight markers, or as
CC nutritional sources or supplements. The proteins may be used to maintain
CC and expand cell population in a totipotential or pluripotential state
CC useful for re-engineering damaged or diseased tissues, transplantation,
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC polynucleotides and proteins are useful for preventing, treating or
CC ameliorating disorders involving aberrant protein expression or
CC biological activity, e.g. haematopoietic disorders, central/peripheral
CC nervous system diseases, mechanical and traumatic disorders, non-healing
CC wounds, immune deficiencies and disorders, infectious diseases caused by
CC viral, bacterial or fungal infection, autoimmune disorders, allergic
CC reactions and conditions, coagulation disorders, or cancer. The
CC polynucleotide sequences of the invention were assembled from ESTs
CC isolated mainly by sequencing by hybridisation, and in some cases,
CC sequences obtained from one or more public databases. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.46e-206	Length:	2008
Score:	2195.00	Matches:	417
Percent Similarity:	99.76%	Conservative:	1
Best Local Similarity:	99.52%	Mismatches:	1

Query Match:	99.50%	Indels:	0
DB:	6	Gaps:	0

US-10-798-532-4 (1-419) x ABQ99383 (1-2008)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 138

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 198

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||
Db        199 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 258

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 318

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 378

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 438

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGCGCCCGA 498

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 558

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 618

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 678

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 738

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 798

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 858

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 918

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        919 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 978

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        979 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1038

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1039 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1098
```

```
Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1099 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1158

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1159 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1218

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1219 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1278

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1279 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1335

RESULT 5
AAI66828
ID   AAI66828 standard; cDNA; 2622 BP.
XX
AC   AAI66828;
XX
DT   07-JAN-2002  (first entry)
XX
DE   Human protein kinase polypeptide 13887 encoding cDNA.
XX
KW   Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KW   cytostatic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1. .266
FT                   /*tag=  a
FT   CDS             267. .1529
FT                   /*tag=  b
FT                   /product= "protein kinase 13887"
FT                   /note= "coding sequence specifically claimed"
FT   3'UTR           1530. .2622
FT                   /*tag=  c
XX
PN   WO200173050-A2.
XX
PD   04-OCT-2001.
XX
PF   23-MAR-2001; 2001WO-US009483.
XX
PR   24-MAR-2000; 2000US-0191846P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Meyers R;
XX
DR   WPI; 2001-611632/70.
DR   P-PSDB; AAG65767.
XX
PT   New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT   useful in diagnosis of cancer or cellular proliferation or
PT   differentiation disorders and to screen for polypeptide modulators useful
PT   to treat such conditions.
XX
PS   Claim 1; Fig 10A-C; 169pp; English.
XX
CC   The invention provides novel human protein kinase polypeptides, 3714,
CC   16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC   protein kinase polypeptides can be expressed by standard recombinant
CC   methodology.3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC   are useful for diagnostic and screening methods to identify subjects (at
CC   risk of) having cancer or cellular proliferation and/or differentiation
CC   disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC   modulators are useful for the treatment of cancer, particularly colon
CC   cancer or cellular proliferation and/or differentiation disorders. Other
CC   disorders associated with 3714, 16742, 23546 or 13887 expression or
CC   activity that can be treated include bone related disorders, inflammatory
CC   disorders, autoimmune diseases, cardiovascular disorders and liver
CC   diseases. The present sequence represents a human protein kinase
CC   polypeptide 13887 encoding cDNA
XX
SQ   Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                5.11e-206      Length:          2622
Score:                    2195.00        Matches:         417
Percent Similarity:       99.76%         Conservative:    1
Best Local Similarity:    99.52%         Mismatches:      1
Query Match:              99.50%         Indels:          0
DB:                       4              Gaps:            0

US-10-798-532-4 (1-419) x AAI66828 (1-2622)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 326

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 386

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        387 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 446

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 506

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 566

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 626

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 686

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 746

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 806

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 866

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 926

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 986

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1046

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1106
```

```
Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1166

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1226

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1286

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1346

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1406

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1407 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1466

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1467 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1523
```

RESULT 6
AAH72714
ID AAH72714 standard; cDNA; 4065 BP.
XX
AC AAH72714;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 3988.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 810-811; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	9.69e-206	Length:	4065
Score:	2195.00	Matches:	417
Percent Similarity:	99.76%	Conservative:	1
Best Local Similarity:	99.52%	Mismatches:	1
Query Match:	99.50%	Indels:	0
DB:	4	Gaps:	0

US-10-798-532-4 (1-419) x AAH72714 (1-4065)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1524
```

RESULT 7
ADL45238
ID ADL45238 standard; DNA; 4065 BP.
XX
AC ADL45238;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human ovarian cancer DNA marker #19128.
XX
KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
OS Homo sapiens.
XX
PN WO200170979-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009126.
XX
PR 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lee J, Lillie J;
XX
DR WPI; 2001-611502/70.
XX
PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX
PS Disclosure; SEQ ID NO 19128; 106pp; English.
XX
CC The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.
XX
SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	9.69e-206	Length:	4065
Score:	2195.00	Matches:	417
Percent Similarity:	99.76%	Conservative:	1
Best Local Similarity:	99.52%	Mismatches:	1
Query Match:	99.50%	Indels:	0
DB:	5	Gaps:	0

US-10-798-532-4 (1-419) x ADL45238 (1-4065)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
                ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1524

RESULT 8
AAS06734
ID   AAS06734 standard; cDNA; 1260 BP.
XX
AC   AAS06734;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Polynucleotide sequence encoding human protein kinase #34.
XX
KW   Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW   metabolic disorder; immune related disease; neurological disorder;
KW   neurodegenerative disorder; inflammatory disorder; infectious disease;
KW   reproductive disorder; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200138503-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US032085.
XX
PR   24-NOV-1999;   99US-0167482P.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman GD,  Whyte D,  Manning G,  Sudarsanam S,  Martinez R;
PI   Flanagan P,  Clary D;
XX
DR   WPI; 2001-343950/36.
DR   P-PSDB; AAU03534.
XX
PT   Nucleic acids encoding human kinase polypeptides, useful for preventing
PT   diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT   neuronal-associated diseases, and microbial infections.
XX
PS   Example 1; Fig 1; 433pp; English.
XX
CC   AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
CC   protein kinases have been identified as members of the tyrosine or
CC   serine/threonine kinase (PTK and STK) families. The polynucleotides
CC   encoding protein kinases and the polypeptides may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate kinase expression. For example, they may be used to treat
CC   cancers (especially cancers of haematopoietic origin), cardiovascular
CC   disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC   immune related diseases (e.g. rheumatoid arthritis), neurological
CC   disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC   Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC   disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC   Additionally, polynucleotides encoding protein kinases may be used for
CC   gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC   polypeptides may be used as antigens in the production of antibodies
CC   against the protein kinases and in assays to identify modulators of
CC   protein kinase expression and activity
XX
SQ   Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              8.59e-206       Length:         1260
Score:                  2188.00         Matches:        416
Percent Similarity:     99.52%          Conservative:   1
Best Local Similarity:  99.28%          Mismatches:     2
Query Match:            99.18%          Indels:         0
DB:                     4               Gaps:           0

US-10-798-532-4 (1-419) x AAS06734 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257
```

```
RESULT 9
AAD18818
ID   AAD18818 standard; cDNA; 1824 BP.
XX
AC   AAD18818;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human kinase (PKIN)-3 cDNA.
XX
KW   Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
KW   cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
KW   AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
KW   atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
KW   myasthenia gravis; cirrhosis; cataract; growth and development disorder;
KW   seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW   lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW   asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
KW   antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             169. .1428
FT                   /*tag= a
FT                   /product= "Human PKIN-3 protein"
XX
PN   WO200181555-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-US012992.
XX
PR   20-APR-2000; 2000US-0199021P.
PR   28-APR-2000; 2000US-0200226P.
PR   05-MAY-2000; 2000US-0202339P.
PR   11-MAY-2000; 2000US-0203505P.
PR   18-MAY-2000; 2000US-0205564P.
PR   26-MAY-2000; 2000US-0207739P.
PR   01-JUN-2000; 2000US-0208795P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Gandhi AR,  Tribouley CM,  Kearney L,  Griffin JA,  Nguyen DB;
PI   Bandman O,  Lu DAM,  Lal P,  Burford N,  Khan FA,  Walia NK,  Yao MG;
PI   Patterson C,  Burrill JD,  Marcus GA,  Zingler KA,  Recipon SA,  Lu Y;
PI   Policky JL,  Thornton M,  Tang YT,  Hafalia A,  Elliott VS,  Baughn MR;
PI   Walsh RT,  Ramkumar J,  Borowsky ML,  Au-Young J,  Hillman JL;
PI   Gururajan R;
XX
DR   WPI; 2001-611740/70.
DR   P-PSDB; AAE11769.
XX
PT   Human kinases and nucleic acids, useful for preventing diagnosing and
PT   treating cancers, inflammation and immune disorders.
XX
PS   Claim 5; Page 152; 166pp; English.
XX
CC   The present invention relates to human kinases (PKIN) and the nucleic
CC   acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
CC   used in the prevention, diagnosis and treatment of diseases cancers,
CC   adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
CC   acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
CC   gout, microbial infections, cardiovascular disease and/or inflammation,
CC   myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
CC   infarction, cataract, growth and development disorder, seizure disorder,
CC   pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
CC   disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
CC   PKIN may be used to treat disorders associated with decreased PKIN
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of PKIN by expressing inactive proteins or to
CC   supplement the patients own production of PKIN. PKIN nucleic acids may be
CC   used to produce the PKIN polypeptide, by inserting the nucleic acids into
CC   a host cell and culturing the cell to express the protein. PKIN nucleic
CC   acid and its complementary sequences may also be used as DNA probes in
CC   diagnostic assays to detect and quantitate the presence of similar
CC   nucleic acid sequences in samples and therefore which patients may be in
CC   need of restorative therapy. The present sequence is human PKIN-3 cDNA
XX
SQ   Sequence 1824 BP; 434 A; 416 C; 472 G; 502 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:              1.47e-205       Length:         1824
Score:                  2188.00         Matches:        416
Percent Similarity:     99.52%          Conservative:   1
Best Local Similarity:  99.28%          Mismatches:     2
Query Match:            99.18%          Indels:         0
DB:                     4               Gaps:           0

US-10-798-532-4 (1-419) x AAD18818 (1-1824)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         169 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 228

Qy          21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         229 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 288

Qy          41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
                ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db         289 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 348

Qy          61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         349 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 408

Qy          81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         409 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 468

Qy         101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         469 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 528

Qy         121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         529 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 588

Qy         141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         589 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 648

Qy         161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         649 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 708

Qy         181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         709 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 768

Qy         201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         769 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 828

Qy         221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         829 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 888

Qy         241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         889 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 948

Qy         261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         949 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1008

Qy         281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1009 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1068

Qy         301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1069 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1128

Qy         321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1129 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1188

Qy         341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1189 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1248

Qy         361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
                ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        1249 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1308

Qy         381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1309 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1368

Qy         401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1369 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1425
```

RESULT 10
ADB52857
ID ADB52857 standard; DNA; 3244 BP.
XX
AC ADB52857;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3399.
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN WO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
DR WPI; 2003-731472/69.
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 3399; 874pp; English.
XX
CC The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity

```
CC   markers for the prediction or identification of the physiological state
CC   of tissue or cell sample that has been exposed to a compound, or to
CC   identify or predict the toxic effects of a compound or an agent. These
CC   may also be used as markers for monitoring toxicity progression or for
CC   drug screening. The present sequence represents a primary rat hepatocyte
CC   toxicity modelling related gene sequence from the present invention.
XX
SQ   Sequence 3244 BP; 800 A; 710 C; 819 G; 915 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.07e-204       Length:         3244
Score:                  2183.00         Matches:        414
Percent Similarity:     99.52%          Conservative:   3
Best Local Similarity:  98.81%          Mismatches:     2
Query Match:            98.96%          Indels:         0
DB:                     10              Gaps:           0

US-10-798-532-4 (1-419) x ADB52857 (1-3244)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 221

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 281

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||| ||||||||||||||||||||||:::||||||||||||||||||||||
Db        282 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 341

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 401

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 461

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 521

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              |||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 581

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 641

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 701

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 761

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 821

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 881

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 941

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        942 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 1001

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 1061

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 AGCCCATTCTTTAGCATTCCTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 1121

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||   ||||||:::||||||
Db       1122 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1181

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1182 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1241

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1242 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1301

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1302 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1361

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1362 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 1418

RESULT 11
ADQ24401
ID   ADQ24401 standard; DNA; 3998 BP.
XX
AC   ADQ24401;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 7221.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 7221; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
```

```
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 3998 BP; 1066 A; 777 C; 885 G; 1255 T; 0 U; 15 Other;

Alignment Scores:
Pred. No.:              1.06e-190       Length:         3998
Score:                  2042.50         Matches:        395
Percent Similarity:     94.76%          Conservative:   3
Best Local Similarity:  94.05%          Mismatches:     21
Query Match:            92.59%          Indels:         1
DB:                     12              Gaps:           1

US-10-798-532-4 (1-419) x ADQ24401 (1-3998)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 219

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 279

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        280 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 339

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||
Db        340 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTAAGCATGTTTCTTACCAGACCG 399

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHis--- 99
              :::                          ||||||||||||||||||||||||||
Db        400 AAGGTCTGTGTTGACCTTAACCGGAGAGTGACTTTGTATGGAGTGTTTACAATCCAGGTT 459

Qy        100 PheSerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSer 119
                     :::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 GTTCTCCAAAGTGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCG 519

Qy        120 GluLeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAla 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 GAATTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCC 579

Qy        140 ArgAspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeu 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 CGAGATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTC 639

Qy        160 LysProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGly 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 AAACCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGA 699

Qy        180 LeuSerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAla 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 CTTAGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCT 759

Qy        200 ProGluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCys 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 CCAGAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGT 819

Qy        220 ThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 879

Qy        240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 939

Qy        260 HisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAsp 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        940 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 999

Qy        280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 1059

Qy        300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCC 1119

Qy        320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGlu 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAA 1179

Qy        340 TyrGluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSer 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 TATGAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCT 1239

Qy        360 LeuLeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAla 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 CTACTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCT 1299

Qy        380 GlyAspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheVal 399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 GGTGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTT 1359

Qy        400 ValAlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1360 GTGGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1419

RESULT 12
AAH70764
ID   AAH70764 standard; cDNA; 529 BP.
XX
AC   AAH70764;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 2038.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer and
PT   for assessing and detecting compounds for treating the cancer.
XX
PS   Claim 1; Page 434; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
```

```
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.8e-68       Length:         529
Score:                  789.00        Matches:        159
Percent Similarity:     95.78%        Conservative:   0
Best Local Similarity:  95.78%        Mismatches:     5
Query Match:            35.77%        Indels:         2
DB:                     4             Gaps:           0

US-10-798-532-4 (1-419) x AAH70764 (1-529)

Qy        220 ThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 85

Qy        240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 145

Qy        260 HisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAsp 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 205

Qy        280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 265

Qy        300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCC 325

Qy        320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGlu 339
              ||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        326 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAA 385

Qy        340 TyrGluAsp-ValValGluAspValLysGlu-GluCysGlnLysTyrGlyProValValS 359
              |||||||| |||||||||||||||||||||| ||||||||||||||||||||    ||||
Db        386 TATGAAGATTGTTGTAGAAGATGTAAAAGAGGGAGTGTCAAAAATATGGACCAGGGGTAT 445

Qy        359 erLeuLeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnA 379
              ||||||||   ||||||||||||||||||||||   ||||||||||||||||||    |
Db        446 CTCTACTTGGTCCAAAGGAAAATCCTGGCAGAGGAACAGTCTTTGTTGAGTATGCAAAGG 505

Qy        379 laGlyAspSerLys 383
              |||||||||||||
Db        506 CTGGGGATTCAAAG 519

RESULT 13
AAA02535
ID   AAA02535 standard; cDNA; 722 BP.
XX
AC   AAA02535;
XX
DT   19-MAY-2000  (first entry)
XX
DE   Human colon cancer cell line polynucleotide sequence SEQ ID NO:2526.
XX
KW   Human; colon cancer; tumour; diagnosis; gene expression product; probe;
KW   detection; cancerous state; metastasis; identification; breast cancer;
KW   oestrogen receptor-positive breast cancer; therapy;
KW   oestrogen receptor-negative breast cancer; lung cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO9958675-A2.
XX
PD   18-NOV-1999.
XX
PF   13-MAY-1999;   99WO-US010602.
XX
PR   14-MAY-1998;   98US-0085426P.
PR   15-MAY-1998;   98US-0085537P.
PR   15-MAY-1998;   98US-0085696P.
PR   21-OCT-1998;   98US-0105234P.
PR   27-OCT-1998;   98US-0105877P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Williams LT,  Escobedo J,  Innis MA,  Garcia PD,  Sudduth-Klinger J;
PI   Reinhard C,  Giese K,  Randazzo F,  Kennedy GC,  Pot D,  Kassam A;
PI   Lamson G,  Drmanac R,  Crkvenjakov R,  Dickson M,  Drmanac S,  Labat I;
PI   Leshkowitz D,  Kita D,  Garcia V,  Jones LW,  Stache-Crain B;
XX
DR   WPI; 2000-126369/11.
XX
PT   Polynucleotide library used to determine cancerous states of mammalian
PT   cells.
XX
PS   Claim 1; Page 1020; 1097pp; English.
XX
CC   AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC   libraries constructed from human colon cancer cell lines. The present
CC   invention also describes a method of detecting differentially expressed
CC   genes correlated with a cancerous state of a mammalian cell, comprising
CC   detecting at least one differentially expressed gene product in a test
CC   sample derived from a cell suspected of being cancerous, where detection
CC   of the differentially expressed gene product is correlated with a
CC   cancerous state of the cell from which the test sample was derived. The
CC   polynucleotides sequences can be used in a method for detecting
CC   differentially expressed genes correlated with a cancerous state of a
CC   mammalian cell. The polynucleotides can also be used as probes for
CC   detecting and mapping related genes. They can be used in diagnosis and
CC   prognosis of diseases and disorders (e.g. identification of pre-
CC   metastatic or metastatic cancerous states, stages of cancer, or
CC   responsiveness of cancer to therapy). This is particularly for breast
CC   cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC   negative breast cancer, lung cancer, and colon cancer
XX
SQ   Sequence 722 BP; 215 A; 129 C; 164 G; 193 T; 0 U; 21 Other;

Alignment Scores:
Pred. No.:              7.32e-66      Length:         722
Score:                  766.50        Matches:        182
Percent Similarity:     84.86%        Conservative:   3
Best Local Similarity:  83.49%        Mismatches:     20
Query Match:            34.75%        Indels:         14
DB:                     3             Gaps:           1

US-10-798-532-4 (1-419) x AAA02535 (1-722)

Qy        152 GluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGlu 171
              ||||||||||||||||||    ||||||   ||||||||||||||||||||||||||||||
Db          1 GAGGGCTATGTCCATGCGGNCCTCAAACNACGTAACATATTGTGGAGTGCAGAGAATGAA 60

Qy        172 CysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspValLysTyr 191
              ||||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db         61 TGTTTTAAACTCATTGACTTTGGACTTANCTTCAAAGAAGGCAATCAGGATGTAAAGTAT 120

Qy        192 IleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAla 211
              ||||||||||||||||||||||||||||   ||||||||||||||||||    |||
Db        121 ATTCAGACAGACGGGTATCGGGCTCCAGAACA-GAATTGCAAAATTGCTTGGCCCANGCT 179

Qy        212 GlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIle 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATT 239

Qy        232 LeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGluTrpLys 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGATCTCAGGAATGGAAG 299
```

```
Qy        252 AlaAsnSer-SerAlaIleIle-AspHisIlePheAlaSerLysAla--ValValAsnAl 270
              ||||||||| |||||||||||| ||||||||||||||||||||:::     ||||||||
Db        300 GCAAACAGTTTCTGCTATTATTTGATCACATATTTGCCAGTAAAAGCAANTGGTGAATGC 359

Qy        270 aAlaIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAsp-ProS 290
              |||||||||||||||||||||    |||||||||||||||||||||||||||||| ||||
Db        360 CGCAATTCCAGCCTATCACCTAANAGACCTTATCAAAAGCATGCTTCATGATGATCCCAA 419

Qy        290 erArg-ArgIleProAlaGluMetAlaLeu-CysSerProPhePheSer-IleProPheA 309
              ||||| ||||||      |||||||||||| |||   |||||||||||| ||||||||||
Db        420 GCAGGAAGAATTTCTNCTGAAATGGCATTGGTGCANCCCATTCTTTAGCNATTCCTTTTG 479

Qy        309 laProHisIleGluAspLeuValMetLeuProThrProVal-LeuArgLeuLeuAsnVal 328
              |||||||||||||||||||    ||||||   |||||||| ||||||||||||||||||
Db        480 CCCCTCATATTGAAGATCTGNTCATGCTTTCCACTCCAGTGGCTAAGACTGCTGAATGTG 539

Qy        329 Leu-AspAspAspTyrLeuGlyAsnGlu--GluGluTyrGluAspValValGluAspVal 347
              |||    ||||||||||||   ||||||  ::::::|||||||||||||||||||||
Db        540 CTGGGNTGATGATTATCTTGAGAATGAAAGAAGGATTATGAAGATGTTGTTGAAGATGNT 599

Qy        348 LysGluGluCysGlnLys---TyrGlyPro 356
              |||          |||      ||||||
Db        600 AAAAGAAGAAGTGGCAAAAATNTTGGACCA 629

RESULT 14
AAH71966
ID   AAH71966 standard; cDNA; 470 BP.
XX
AC   AAH71966;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 3240.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer and
PT   for assessing and detecting compounds for treating the cancer.
XX
PS   Claim 1; Page 636; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.96e-62        Length:          470
Score:                    729.00          Matches:         147
Percent Similarity:       96.71%          Conservative:    0
Best Local Similarity:    96.71%          Mismatches:      3
Query Match:              33.05%          Indels:          2
DB:                       4               Gaps:            0

US-10-798-532-4 (1-419) x AAH71966 (1-470)

Qy        220 ThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 77

Qy        240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 137

Qy        260 HisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAsp 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 197

Qy        280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 257

Qy        300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCC 317

Qy        320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGlu 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAA 377

Qy        340 TyrGluAsp-ValValGluAspValLysGlu-GluCysGlnLysTyrGlyProValValS 359
              |||||||| |||||||||||||||||||| ||||||||||||||||||||    ||||
Db        378 TATGAAGATTGTTGTAGAAGATGTAAAAGAGGGAGTGTCAAAAATATGGACCAGGGGTAT 437

Qy        359 erLeuLeuValProLysGluAsnProGlyArg 369
              ||||||||   ||||||||||||||||||||
Db        438 CTCTACTTGGTCCAAAGGAAAATCCTGGCAGA 469

RESULT 15
AAH69101
ID   AAH69101 standard; cDNA; 461 BP.
XX
AC   AAH69101;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 375.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US033312.
XX
PR   08-DEC-1999;   99US-0169681P.
PR   21-DEC-1999;   99US-0171350P.
PR   14-MAR-2000; 2000US-0189315P.
PR   12-MAY-2000; 2000US-0203791P.
PR   09-JUN-2000; 2000US-0210600P.
PR   21-JUL-2000; 2000US-0220114P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
```

DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 168-169; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 461 BP; 134 A; 91 C; 109 G; 127 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.87e-53	Length:	461
Score:	636.00	Matches:	138
Percent Similarity:	93.33%	Conservative:	2
Best Local Similarity:	92.00%	Mismatches:	5
Query Match:	28.83%	Indels:	6
DB:	4	Gaps:	0

US-10-798-532-4 (1-419) x AAH69101 (1-461)

```
Qy        220 ThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACCTCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 81

Qy        240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 AAACTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGAT 141

Qy        260 HisIlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAsp 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 CACATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGAC 201

Qy        280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 CTTATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTG 261

Qy        300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
              |||||||||||||||||||||||||||||||||||    ||||||||||||||||||||
Db        262 TGCAGCCCATTCTTTAGCATTCCTTTTGCCCCTCATAT-GAAGATCTGGTCATGCTTCCC 320

Qy        320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGlu 339
              |||||||||||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db        321 ACTCCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAA 380

Qy        340 TyrGluAsp--ValValGluAspValLysGlu--GluCysGlnLysTyrGlyPro-ValV 358
              |||||||||  ::::::||||||||||||||||     ||||||||||||||||| ||||
Db        381 TATGAAGATTGTTGTTAGAAGATGTAAAAGAGGGAGGTGTCAAAAATATGGACCAGGTGG 440

Qy        358 alSerLeuLeuValProLys 364
              ||||||||||||      |||
Db        441 TATCTCTACTTTGTTCCAAA 460
```

Search completed: November 29, 2004, 11:42:36
Job time : 520 secs

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 26 Seconds
(without alignments)
1550.570 Million cell updates/sec

Title: US-10-798-532-4
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAFG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1052	47.7	200	2	I48615	gene KIS protein -
2	248.5	11.3	948	2	T24445	hypothetical prote
3	240.5	10.9	720	2	A96807	hypothetical prote
4	240.5	10.9	826	2	JC5153	mitogen-activated
5	240	10.9	388	2	S48879	protein kinase SMK
6	238.5	10.8	387	2	T37758	protein kinase skp
7	234.5	10.6	433	2	S37790	probable serine/th
8	233	10.6	369	2	A56492	protein kinase ERK
9	232.5	10.5	367	2	S68680	stress-activated p
10	231.5	10.5	354	2	T45138	protein kinase skp
11	231	10.5	435	2	S64950	protein kinase HOG
12	230	10.4	298	1	A44878	protein kinase (EC
13	230	10.4	298	2	A41227	protein kinase (EC
14	230	10.4	326	2	S23387	protein kinase (EC
15	230	10.4	373	2	T13024	probable protein k
16	230	10.4	376	2	S40470	mitogen-activated
17	230	10.4	376	2	S40471	mitogen-activated
18	230	10.4	427	2	JC5693	stress-activated p
19	228	10.3	423	2	S43968	p54-alpha-2 stress
20	227	10.3	423	2	S43967	p54-alpha stress-a
21	227	10.3	424	2	S71102	protein kinase JNK
22	227	10.3	529	1	S71774	calcium-dependent
23	226	10.2	406	1	KIHUCT	phosphorylase kina
24	225.5	10.2	301	1	S19209	protein kinase (EC
25	225.5	10.2	416	2	A48249	pre-mRNA splicing
26	225	10.2	371	2	S60121	mitogen-activated
27	224.5	10.2	335	2	T23050	hypothetical prote
28	224.5	10.2	342	2	T21098	protein kinase (EC
29	224.5	10.2	371	2	T14915	mitogen-activated
30	224.5	10.2	469	1	S17656	Ca2+/calmodulin-de
31	223.5	10.1	367	2	JC5252	mitogen-activated
32	223.5	10.1	426	2	S43969	p54-beta stress-ac
33	223.5	10.1	464	2	S71104	protein kinase JNK
34	223.5	10.1	678	2	T43539	spindle checkpoint
35	223	10.1	346	1	I48157	protein kinase (EC
36	223	10.1	406	2	D84898	probable mitogen-a
37	223	10.1	427	2	JC5694	stress-activated p
38	223	10.1	575	2	JC7794	lammer kinase homo
39	223	10.1	690	2	T38052	probable protein k
40	222.5	10.1	301	1	A48041	protein kinase (EC
41	222.5	10.1	424	2	A55480	c-Jun amino-termin
42	221	10.0	358	2	S23383	protein kinase (EC
43	220.5	10.0	382	2	JC5531	c-Jun amino-termin
44	220.5	10.0	602	2	S60052	calcium-dependent
45	220	10.0	371	2	T09622	protein kinase MMK

ALIGNMENTS

RESULT 1
I48615
gene KIS protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I48615
R;Maucuer, A.; Camonis, J.H.; Sobel, A.
Proc. Natl. Acad. Sci. U.S.A. 92, 3100-3104, 1995
A;Title: Stathmin interaction with a putative kinase and coiled-coil-forming protein dome
A;Reference number: I48282; MUID:95241452; PMID:7724523
A;Accession: I48615
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-200 <RES>
A;Cross-references: UNIPROT:P97343; EMBL:X82320; NID:g791078; PIDN:CAA57763.1; PID:g79107
C;Genetics:
A;Gene: KIS

Query Match 47.7%; Score 1052; DB 2; Length 200;
Best Local Similarity 99.0%; Pred. No. 8.1e-54;
Matches 198; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy        122 LLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db          1 LLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAEDECFKLIDFGLS 60

Qy        182 FKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 FKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL 120

Qy        242 KHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALCS 301
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        121 KHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALCS 180

Qy        302 PFFSIPFAPHIEDLVMLPTP 321
              ||||||||||||||||||||
Db        181 PFFSIPFAPHIEDLVMLPTP 200
```

RESULT 2
T24445
hypothetical protein T04C10.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24445
R;Burton, J.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19891
A;Accession: T24445
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-948 <WIL>

A;Cross-references: UNIPROT:Q8WQL7; EMBL:Z69885; PIDN:CAA93756.1; GSPDB:GN00028; CESP:T0
A;Experimental source: clone T04C10
C;Genetics:
A;Gene: CESP:T04C10.1
A;Map position: X
A;Introns: 64/3; 181/3; 214/3; 238/1; 272/3; 331/2; 391/2; 457/1; 726/3; 830/2; 923/3

Query Match 11.3%; Score 248.5; DB 2; Length 948;
Best Local Similarity 25.7%; Pred. No. 5.6e-07;
Matches 104; Conservative 57; Mismatches 113; Indels 131; Gaps 17;

```
Qy         88 NIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALA 147
Db        455 NIVTLKGHF-VHRA----HLCLVFELLSYNLYDLLKNTSFRGVSLNLARKFAQQLGKTLL 509

Qy        148 FLHHE--GYVHADLKPRNILW-SAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAEL 204
Db        510 FLSSPELSIIHCDLKPENVLLVNAKRSQIRVIDFGSSCQTGHRIYQYIQSRFYRSPEVLL 569

Qy        205 QNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSG----------MKL---------- 241
Db        570 -------GIAYDTK----IDMWSLGCILVEMHTGEPLFAGSSEVDQMMKIVEVLGMPPKE 618

Qy        242 ----------------------------KHTVRSQEWKANSSAIIDHIFASKAVVNAAI 272
Db        619 MLDIGPKTHKYFDKTEDGIYYCKKTRDGYRHT-----YKAPGARKLHEILG----VTSGG 669

Qy        273 PA--------------YHLRDLIKSMLHDDPSRRIPAEMALCSPFFSI--PFAPHIEDLV 316
Db        670 PGGRRLGEPGHSVEDYSKFKDLIKRMLQFDPKQRI-------SPYYVVRHPFLKQKEERV 722

Qy        317 MLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKY--GPVVSLLVPKENPGRGQVFV 374
Db        723 PSQPPVSH--------------------SNLQQQQQLYIQQPSSQMSQVMESPSVGSVYV 762

Qy        375 E----YANAGDSKAAQKLLTGRMFDGKF--VVATFYPLSAYKRGY 413
Db        763 EDNGMYRQAPGSSANPISVTSSFDEGDAMEVDAGRRRFSAHQQNY 807
```

RESULT 3
A96807
hypothetical protein T32E8.5 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A96807
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96807
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-720 <STO>
A;Cross-references: UNIPROT:Q9CA22; GB:AE005173; NID:g6437532; PIDN:AAF08564.1; GSPDB:GN
C;Genetics:
A;Gene: T32E8.5
A;Map position: 1

Query Match 10.9%; Score 240.5; DB 2; Length 720;
Best Local Similarity 25.7%; Pred. No. 1.2e-06;
Matches 105; Conservative 57; Mismatches 146; Indels 101; Gaps 17;

```
Qy         11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPGTTGAAASAAE 70
Db        331 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 383

Qy         71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
Db        384 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 436

Qy        121 ---LLLYS----SHQGCSMW-------------MIQHCARDVLEALAFLHHEGYVHADLK 160
Db        437 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 496

Qy        161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
Db        497 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 551

Qy        215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272
Db        552 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 607

Qy        273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH---------------------- 311
Db        608 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 663

Qy        312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351
Db        664 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 706
```

RESULT 4
JC5153
mitogen-activated protein kinase (EC 2.7.1.-) - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: JC5153
R;Doerig, C.M.; Parzy, D.; Langsley, G.; Horrocks, P.; Carter, R.; Doerig, C.D.
Gene 177, 1-6, 1996
A;Title: A MAP kinase homologue from the human malaria parasite, Plasmodium falciparum.
A;Reference number: JC5153; MUID:97080491; PMID:8921836
A;Accession: JC5153
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-826 <DOE>
A;Cross-references: UNIPROT:Q94656; EMBL:X82646
C;Keywords: phosphotransferase

Query Match 10.9%; Score 240.5; DB 2; Length 826;
Best Local Similarity 25.1%; Pred. No. 1.4e-06;
Matches 103; Conservative 57; Mismatches 145; Indels 105; Gaps 19;

```
Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPP-----GALRQFLPPGTTGAAASAAEYGFRKER 77
Db         23 YDILKKVGKGAYGVVFKGRCKKNKNIVAVKKIFGAFQN----------CTDAQRTFR-EI 71

Qy         78 AALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSEL----LLYSSHQGCSMW 133
Db         72 IFLYELNGHDNIIKLMDVIK---AKNDNDIYLIFDFMETDLHEVIKADLLEEIHK----- 123

Qy        134 MIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEGN 186
Db        124 --KYIIYQLLRALKYIHSGGLLHRDIKPSNIL--VNSECHIKVADFGLARSISTHVNENK 179

Qy        187 QDV--KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL--- 241
Db        180 VPILTDYVATRWYRAPEI----------LLGSTHYTEDVDMWSLGCIMGELLCGKPLFTG 229

Qy        242 KHTVRSQE------WKANSSAIID--HIFASKAVVN-AAIPAYHLR-----------DLI 281
Db        230 NSTMNQLEKIIQVIGKPNKKDIEDIRSPFAEKIISSFVDLKKKNLKDICYKASNESLDLL 289

Qy        282 KSMLHDDPSRRIPAEMAL----CSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNE 337
Db        290 EKLLQFNPSKRISAENALKHKYVEEFHSIIDEPTCRHIITIP----------------- 331

Qy        338 EEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQK 387
```

```
Db        332 -----INDNTKYRVNFYRNVVYFVIMRRNKFHSNVL----NQGESKKEEK 372


RESULT 5
S48879
protein kinase SMK1 (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YP9499.10; protein YPR054w
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C;Accession: S48879; S54076
R;Krisak, L.; Strich, R.; Winters, R.S.; Hall, J.P.; Mallory, M.J.; Kreitzer, D.; Tuan,
Genes Dev. 8, 2151-2161, 1994
A;Title: SMK1, a developmentally regulated MAP kinase, is required for spore wall assemb
A;Reference number: S48879; MUID:95047362; PMID:7958885
A;Accession: S48879
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-388 <KRI>
A;Cross-references: UNIPROT:P41808; EMBL:L35047; NID:g538516; PIDN:AAB59325.1; PID:g5385
R;Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54059
A;Accession: S54076
A;Molecule type: DNA
A;Residues: 1-388 <BAD>
A;Cross-references: EMBL:Z49219; NID:g805025; PIDN:CAA89172.1; PID:g805035; MIPS:YPR054w
A;Experimental source: strain AB972
C;Genetics:
A;Gene: SGD:SMK1
A;Cross-references: SGD:S0006258; MIPS:YPR054w
A;Map position: 16R
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase; signal transduction
F;36-306/Domain: protein kinase homology <KIN>
F;44-52/Region: protein kinase ATP-binding motif

  Query Match             10.9%;  Score 240;  DB 2;  Length 388;
  Best Local Similarity   25.8%;  Pred. No. 6.6e-07;
  Matches  103;  Conservative   61;  Mismatches  133;  Indels  102;  Gaps   19;

Qy         21 RLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPGTTGAAASAAEYG-------- 72
              |   | | ||:    :::        ||     : |||  |  |    |    |
Db         10 RAINVASNLGAPQQRTIFAKERISIPGY--YEIIQFLGKGAYGTVCSVKFKGRSPAARIA 67

Qy         73 -------FRKE---RAALEQL------QGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDV 116
                     | ||   : |: :|      :||:||| |  :  :  ||     |   ||:|
Db         68 VKKISNIFNKEILLKRAIRELKFMNFPKGHKNIVNLIDLEIVTSSPYDGLYC-YQELIDY 126

Qy        117 SVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLI 176
               ::: :::|| |   | : |::    :|  | ::|    :| |||| ||| :  | | |:
Db        127 DLAK-VIHSSVQ-LSEFHIKYFLYQILCGLKYIHSADVIHRDLKPGNILCTL-NGCLKIC 183

Qy        177 DFGLS------FKEGNQDVK-----YIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDL 225
              ||||:      | : :  |:     |: |  |||||           | |:   : :||:
Db        184 DFGLARGIHAGFFKCHSTVQPHITNYVATRWYRAPEL----------LLSNQPYSKSVDI 233

Qy        226 WSLGIILLE------MPSGMKLKHTV------------------RSQEW---KANSSAI 257
              |::| || |      :| |    | :                  : |   | ::: :
Db        234 WAVGCILAEFYARKPVFMGRDSMHQIFEIIKVLGTPDKDILIKFGTIKAWNLGKNSNNPV 293

Qy        258 IDHI-------FASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAP 310
                 |       |||   :|          ||:|:|| | : |:  | |:  || :    |
Db        294 YKKIPWSNIFPFASHEAIN----------LIESLLHWDSTHRLNVEQAISHPFLNEVRKP 343

Qy        311 HIEDLVMLPTPVLRLLNVLDDDYLGNEEEYED-VVEDVK 348
              :: | |  |     :   :  | :  :  | :||:||
Db        344 D-DEPVCLQGP----FDFTYESELNSMSKLRDYLVEEVK 377


RESULT 6
T37758
protein kinase skp1p - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37758
R;Wedler, H.; Wambutt, R.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1998
A;Reference number: Z21744
A;Accession: T37758
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-387 <WED>
A;Cross-references: UNIPROT:Q10452; EMBL:AL035064; PIDN:CAA22609.1; GSPDB:GN00066; SPDB:S
A;Experimental source: strain 972h-; cosmid c1687
C;Genetics:
A;Gene: SPDB:SPAC1687.15
A;Map position: 1
A;Introns: 7/1; 70/3
C;Superfamily: kinase-related transforming protein; protein kinase homology

  Query Match             10.8%;  Score 238.5;  DB 2;  Length 387;
  Best Local Similarity   24.6%;  Pred. No. 8.1e-07;
  Matches   78;  Conservative   48;  Mismatches  112;  Indels   79;  Gaps   10;

Qy         29 LGSGSSASVYRVRCCGNPGSPPGALRQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRN 88
              :||||   | :|    :      |::: |           :  |:     : ::  | |
Db         38 VGSGSFGVVMQVHLIESDSK--AAIKRVL----------QDKRFKNRELQIMRIMKHPN 84

Qy         89 IVTLYGVF-------------TIHFSPNV---PSRCLLLELLDVSVSELLLYSSHQGCS 131
              || |   :               : | |      ||    : | : : |: ||
Db         85 IVDLIAYYYTTGDNSDEVYLNLVLEFMPETIYRASRLYTRQKLSMPMLEVKLY------- 137

Qy        132 MWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLS--FKEGNQDV 189
                        :| :||::|  |  | |:||:|:|   ||    || ||| :     |  :|
Db        138 -------IYQLLRSLAYIHASGICHRDIKPQNLLLDPENGILKLCDFGSAKILVAGEPNV 190

Qy        190 KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSG---------- 238
               || :  |||||           :   |: | |:|:|| | :: |:  |
Db        191 SYICSRYYRAPEL----------IFGATDYTHAIDIWSTGCVMAELMLGHPLFPGESGID 240

Qy        239 -----MKLKHTVRSQEWKANSSAIIDHIFAS------KAVVNAAIPAYHLRDLIKSMLHD 287
                   :|:  |   :: |  :   ::| |           | : ::|   | ||:  ||
Db        241 QLVEIIKILGTPSREQIKTMNPNYMEHRFPQIRPQPLSRVFSRSVPLDAL-DLLSKMLQY 299

Qy        288 DPSRRIPAEMALCSPFF 304
               |: |: |  |:| |||
Db        300 TPTDRLTAAEAMCHPFF 316


RESULT 7
S37790
probable serine/threonine-specific protein kinase (EC 2.7.1.-) YKL161c - yeast (Saccharom
N;Alternate names: protein YKL615
C;Species: Saccharomyces cerevisiae
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 16-Aug-2004
C;Accession: S37790; S37991; S44567
R;Vandenbol, M.; Bolle, P.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the EMBL Data Library, September 1993
A;Description: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4 loc
A;Reference number: S37786
A;Accession: S37790
A;Molecule type: DNA
A;Residues: 1-433 <VAN>
A;Cross-references: UNIPROT:P36005; EMBL:Z26877; NID:g407482; PIDN:CAA81493.1; PID:g40748
A;Experimental source: strain S288C
R;Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37976
A;Accession: S37991
A;Molecule type: DNA
A;Residues: 1-433 <VA2>
A;Cross-references: EMBL:Z28161; NID:g486280; PIDN:CAA82003.1; PID:g486281; MIPS:YKL161c
A;Experimental source: strain S288C
R;Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
Yeast 10, 35-40, 1994
A;Title: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4 loci of c
```

```
A;Reference number: S44563
A;Accession: S44567
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-433 <VA3>
A;Cross-references: EMBL:Z26877; NID:g407482; PIDN:CAA81493.1; PID:g407487
A;Experimental source: strain S288C
C;Genetics:
A;Cross-references: SGD:S0001644
A;Map position: 11L
A;Note: YKL161c
C;Superfamily: protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;21-287/Domain: protein kinase homology <KIN>
F;153/Active site: Asp #status predicted

  Query Match             10.6%;  Score 234.5;  DB 2;  Length 433;
  Best Local Similarity   29.5%;  Pred. No. 1.5e-06;
  Matches   89;  Conservative   45;  Mismatches   97;  Indels   71;  Gaps   16;

Qy         75 KERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLL--ELLDVSVSELLLYSSHQGCSM 132
Db         71 RELKLLRHLRGHPNIVWLFDT-DIVFYPNGALNGVYLYEELMECDLSQII--RSEQRLED 127

Qy        133 WMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEG 185
Db        128 AHFQSFIYQILCALKYIHSANVLHCDLKPKNLL--VNSDCQLKICNFGLSCSYSENHKVN 185

Qy        186 NQDVK-YIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLE------MFSG 238
Db        186 DGFIKGYITSIWYKAPEI----------LLNYQECTKAVDIWSTGCILAELLGRKPMFEG 235

Qy        239 M----KLKHTVR---------SQEWKANSSAIIDHIF--------ASKAVVNAAIPAYHL 277
Db        236 KDYVDHLNHILQILGTPPEETLQE--IASQKVYNYIFQFGNIPGRSFESILPGANP--EA 291

Qy        278 RDLIKSMLHDDPSRRIPAEMALCSPFFSI------PFA---------PHIEDLVMLPTPV 322
Db        292 LELLKKMLEFDPKKRITVEDALEHPYLSMWHDIDEEFSCQKTFRFEFEHIESMAELGNEV 351

Qy        323 LR 324
Db        352 IK 353


RESULT 8
A56492
protein kinase ERK2 (EC 2.7.1.-) - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56492
R;Segall, J.E.; Kuspa, A.; Shaulsky, G.; Ecke, M.; Maeda, M.; Gaskins, C.; Firtel, R.A.;
J. Cell Biol. 128, 405-413, 1995
A;Title: A MAP kinase necessary for receptor-mediated activation of adenylyl cyclase in
A;Reference number: A56492; MUID:95146547; PMID:7844154
A;Accession: A56492
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-369 <SEG>
A;Cross-references: UNIPROT:Q7M445; GB:L33043
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;12-273/Domain: protein kinase homology <KIN>
F;20-28/Region: protein kinase ATP-binding motif

  Query Match             10.6%;  Score 233;  DB 2;  Length 369;
  Best Local Similarity   26.9%;  Pred. No. 1.6e-06;
  Matches   91;  Conservative   52;  Mismatches  109;  Indels   86;  Gaps   17;

Qy         66 ASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSEL---- 121
Db         52 ATDAQRTFR-EIMFLQELHGHENIIKLLNVIK---ADNDRDIYLVFEHMETDLHAVIRAK 107

Qy        122 LLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECF-KLIDFGL 180
Db        108 ILEEIHK-------QYTIYQLLKALKYMHSANVLHRDIKPSNLL--LNSECLVKVADFGL 158

Qy        181 --------SFKEGNQDV-KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGII 231
Db        159 ARSITSLESIAEANPVLTEYVATRWYRAPEI----------LLGSTKYTKGVDMWSIGCI 208

Qy        232 LLE------MFSG----------MKLKHTVRSQEWKANSSAIIDHIFASKAVVN------ 269
Db        209 LGELLGEKAMFPGNSTMNQLDLIIEVTGRPSAEDIEAIKSPFAGTMLESLPPSNPRSLSD 268

Qy        270 ----AAIPAYHLRDLIKSMLHDDPSRRIPAEMALCSPF---FSIPF-APHIEDLVMLPT- 320
Db        269 MYPSASVDAL---DLLKKLSQFNPDKRITAEEALAHPFVTQFHNPAEEPHFDRIIKISID 325

Qy        321 -----PVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQK 353
Db        326 DGQKFPIAEYRNRL----------YNDIIKKKKKEERKK 353


RESULT 9
S68680
stress-activated protein kinase (EC 2.7.1.-) 3 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S68680
R;Mertens, S.; Craxton, M.; Goedert, M.
FEBS Lett. 383, 273-276, 1996
A;Title: SAP kinase-3, a new member of the family of mammalian stress-activated protein k
A;Reference number: S68680; MUID:96198618; PMID:8925912
A;Accession: S68680
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-367 <MER>
A;Cross-references: UNIPROT:Q63538; EMBL:X96488; NID:g1262402; PIDN:CAA65342.1; PID:g1262
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;25-280/Domain: protein kinase homology <KIN>
F;33-41/Region: protein kinase ATP-binding motif

  Query Match             10.5%;  Score 232.5;  DB 2;  Length 367;
  Best Local Similarity   22.6%;  Pred. No. 1.7e-06;
  Matches   95;  Conservative   68;  Mismatches  127;  Indels  131;  Gaps   19;

Qy         10 AEPPRFLEAFGR------LWQVQS------RLGSGSSASVYRVRCCGNPGSPPG---ALR 54
Db          2 SSPPPARKGFYRQEVTKTAWEVRAVYQDLQPVGSGAYGAV-----CSAVDSRTGNKVAIK 56

Qy         55 QFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRC------ 108
Db         57 KLYRPFQSELFAKRA----YRELRLLKHMR-HENVIGLLDVFT-------PDETLDDFTD 104

Qy        109 --LLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILW 166
Db        105 FYLVMPFMGTDLGKLM---KHETLSEDRIQFLVYQMLKGLKYIHAAGVIHRDLKPGNL-- 159

Qy        167 SAENEC-FKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDL 225
Db        160 AVNEDCELKILDFGLARQADSEMTGYVVTRWYRAPEVIL----------NWMRYTQTVDI 209

Qy        226 WSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNA---------------- 270
Db        210 WSVGCIMAEMITGKIL--------FKGN-----DHLDQLKEIMKVTGTPPPEFVQKLQSA 256

Qy        271 -------AIPAYHLRD--------------LIKSMLHDDPSRRIPAEMALCSPFFSIPFA 309
Db        257 EAKNYMEGLPELEKKDFASVLTNASPQAVNLLEKMLVLDAEQRVTAAEALAHPYFE---- 312

Qy        310 PHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVK---EECQK--YGPVVSLLVPK 364
Db        313 ----------------SLRDTEDEPKAQKYDDSFDDVDRTLEEWKRVTYKEVLSFKPPR 355
```

```
Qy        365 E 365
              :
Db        356 Q 356
```

RESULT 10
T45138
protein kinase skp1 [imported] - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T45138
R;Plyte, S.E.; Feoktistova, A.; Burke, J.D.; Woodgett, J.R.; Gould, K.L.
Mol. Cell. Biol. 16, 179-191, 1996
A;Title: Schizosaccharomyces pombe skp1+ encodes a protein kinase related to mammalian g
A;Reference number: Z22927; MUID:96104567; PMID:8524294
A;Accession: T45138
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-354 <PLY>
A;Cross-references: EMBL:L29449; NID:g1914882; PIDN:AAB51081.1; PID:g1914883
C;Genetics:
A;Gene: skp1
A;Introns: 6/3
C;Function:
A;Pathway: cells w
A;Note: cells with deletions in skp1 are sensitive to heat shock and exhibit defects in
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: phosphoprotein; phosphotransferase
F;335/Binding site: phosphate (Ser) (covalent) #status experimental

  Query Match 10.5%; Score 231.5; DB 2; Length 354;
  Best Local Similarity 24.3%; Pred. No. 1.8e-06;
  Matches 77; Conservative 48; Mismatches 113; Indels 79; Gaps 10;

```
Qy         29 LGSGSSASVYRVRCCGNPGSPPGALRQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRN 88
              :||||    | :|    :       |::: |          :  |:     : ::    |
Db         38 VGSGSFGVVMQVHLIESDSK--AAIKRVL----------QDKRFKNRELQIMRIMNDPN 84

Qy         89 IVTLYGVF-------------TIHFSPNV---PSRCLLLELLDVSVSELLLYSSHQGCS 131
              || |   :             : | |      ||    : | : : |: ||
Db         85 IVDLIAYYYTTGDNSDEVYLNLVLEFMPETIYRASRLYTRQKLSMPMLEVKLY------- 137

Qy        132 MWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLS--FKEGNQDV 189
                    :| :||::|  |  | |:||:|:|   ||   || ||| :     |  :|
Db        138 -------IYQLLRSLAYIHASGICHRDIKPQNLLLDPENGILKLCDFGSAKILVAGEPNV 190

Qy        190 KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSG---------- 238
               || :  |||||           :   |: | |:|:|| | :: |:  |
Db        191 SYICSRYYRAPEL----------IFGATDYTHAIDIWSTGCVMAELMLGHPLFPGESGID 240

Qy        239 -----MKLKHTVRSQEWKANSSAIIDHIFAS------KAVVNAAIPAYHLRDLIKSMLHD 287
                  :|:  |   :: |  :   ::| |          | : ::|   | ||:  ||
Db        241 QLVEIIKILGTPSREQIKTMNPNYMEHRFPQIRPQPLSRVFSRSVPLDAL-DLLSKMLQY 299

Qy        288 DPSRRIPAEMALCSPFF 304
               |: |: |  |:| |||
Db        300 TPTDRLTAAEAMCHPFF 316
```

RESULT 11
S64950
protein kinase HOG1 (EC 2.7.1.-), mitogen-activated - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein L2931; protein YLR113w
C;Species: Saccharomyces cerevisiae
C;Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 16-Aug-2004
C;Accession: S64950; S42145; S69400
R;Verhasselt, P.; Voet, M.; Volckaert, G.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64943
A;Accession: S64950
A;Molecule type: DNA
A;Residues: 1-435 <VER>
A;Cross-references: UNIPROT:P32485; EMBL:Z73285; NID:g1360507; PIDN:CAA97680.1; PID:e245
A;Experimental source: strain S288C
R;Brewster, J.L.; de Valoir, T.; Dwyer, N.D.; Winter, E.; Gustin, M.C.
Science 259, 1760-1763, 1993
A;Title: An osmosensing signal transduction pathway in yeast.
A;Reference number: S42145; MUID:93206121; PMID:7681220
A;Accession: S42145
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-408,'GQRSCSCK' <BRE>
A;Cross-references: EMBL:L06279; NID:g295610; PIDN:AAA34680.1; PID:g295611
R;Verhasselt, P.; Volckaert, G.
submitted to the EMBL Data Library, September 1995
A;Reference number: S69393
A;Accession: S69400
A;Molecule type: DNA
A;Residues: 1-435 <VEW>
A;Cross-references: EMBL:X89514; NID:g1297019; PIDN:CAA61691.1; PID:e198746; PID:g129702
C;Genetics:
A;Gene: SGD:HOG1; SSK3
A;Cross-references: SGD:S0004103; MIPS:YLR113w
A;Map position: 12R
C;Superfamily: protein kinase homology
C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase; signal tra
F;21-271/Domain: protein kinase homology <KIN>
F;29-37/Region: protein kinase ATP-binding motif
F;144/Active site: Asp #status predicted

  Query Match 10.5%; Score 231; DB 2; Length 435;
  Best Local Similarity 25.5%; Pred. No. 2.5e-06;
  Matches 97; Conservative 54; Mismatches 140; Indels 90; Gaps 18;

```
Qy         17 EAFGRLWQVQSR------LGSGSSASVYRVRCCGNP--GSPPGALRQFLPPGTTGAAASA 68
              : || :::: :|      :| |:   |    |       | | |::: : | :|   |
Db         11 QIFGTVFEITNRYNDLNPVGMGAFGLV----CSATDTLTSQPVAIKKIMKPFSTAVLAKR 66

Qy         69 AEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQ 128
                 :|   |: |: | |:: |  :|    || :    : ||    :  ||     :
Db         67 T----YRELKLLKHLR-HENLICLQDIF---LSP-LEDIYFVTELQGTDLHRLL---QTR 114

Qy        129 GCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQD 188
                 :|:    :|  | ::|  | :| |||| ||| : ||   |: ||||:  :  |
Db        115 PLEKQFVQYFLYQILRGLKYVHSAGVIHRDLKPSNILIN-ENCDLKICDFGLARIQDPQM 173

Qy        189 VKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL----KHT 244
               |: |  |||||  |           : | |    ||:|| | |  ||  | |      |
Db        174 TGYVSTRYYRAPEIML------TWQKYDVE----VDIWSAGCIFAEMIEGKPLFPGKDH- 222

Qy        245 VRSQEWKANSSAIIDHIFAS--KAVVNAAIPAYHLR------------------------ 278
                    :  :||  :  |  | |:|       |:
Db        223 -------VHQFSIITDLLGSPPKDVINTICSENTLKFVTSLPHRDPIPFSERFKTVEPDA 275

Qy        279 -DLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVM-----------LPTPVLRL 325
               ||:: ||  || :|| |  ||  | :| |:    :: |              ||    |:
Db        276 VDLLEKMLVFDPKKRITAADALAHP-YSAPYHDPTDEPVADAKFDWHFNDADLPVDTWRV 334

Qy        326 L---NVLDDDYLGNEEEYEDV 343
              :   :||   :|  :   |:
Db        335 MMYSEILDFHKIGGSDGQIDI 355
```

RESULT 12
A44878
protein kinase (EC 2.7.1.37) cdk2 [validated] - goldfish
C;Species: Carassius auratus (goldfish)
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: A44878
R;Hirai, T.; Yamashita, M.; Yoshikuni, M.; Tokumoto, T.; Kajiura, H.; Sakai, N.; Nagahama
Dev. Biol. 152, 113-120, 1992
A;Title: Isolation and characterization of goldfish cdk2, a cognate variant of the cell c
A;Reference number: A44878; MUID:92331802; PMID:1339336
A;Accession: A44878
A;Status: preliminary
A;Molecule type: mRNA

```
A;Residues: 1-298 <HIR>
A;Cross-references: UNIPROT:P43450; GB:S40289; NID:g251619; PIDN:AAB22550.1; PID:g251620
A;Experimental source: oocyte
A;Note: sequence extracted from NCBI backbone (NCBIN:108782, NCBIP:108783)
C;Genetics:
A;Gene: cdk2
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine
F;2-255/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

  Query Match            10.4%;  Score 230;  DB 1;  Length 298;
  Best Local Similarity   28.4%;  Pred. No. 1.9e-06;
  Matches   98;  Conservative   49;  Mismatches  116;  Indels   82;  Gaps   17;

Qy         16 LEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQF-LPPGTTGAAASAAEYGFR 74
              :|:|    |   ::| |:   ||: :          ||::  |   | |  ::|
Db          1 MESF----QKVEKIGEGTYGVVYKAK--NKVTGETVALKKIRLDTETEGVPSTAI----- 49

Qy         75 KERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWM 134
              :| : |::|  | ||| |: |  ||    :    |: | |   :  :  |:  | |: :
Db         50 REISLLKEL-NHPNIVKLHDV--IHTENKL---YLVFEFLHQDLKRFMDSSTVTGISLPL 103

Qy        135 IQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY--- 191
              ::     :|: ||| |    :| ||||:|:| :|: |  || ||||:   |      |
Db        104 VKSYLFQLLQGLAFCHSHRVLHRDLKPQNLLINAQGE-IKLADFGLARAFGVPVRTYTHE 162

Qy        192 IQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWK 251
              : |  |||||  |  |            ::|||:|||| |  || :        |   :
Db        163 VVTLWYRAPEI-LLGC---------KYYSTAVDIWSLGCIFAEMIT--------RKALFP 204

Qy        252 ANSSAIIDHIFASKAVVN----------AAIPAYHL--------------------RDL 280
               :|   || :|    :           ::| |                       |||
Db        205 GDSE--IDQLFRIFRTLGTPDESIWPGVTSMPDYKPSFPKWARQDLSKVVPPLDEDGRDL 262

Qy        281 IKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRL 325
              :  ||  ||::|| |: ||  ||        |: | | | |||
Db        263 LGQMLIYDPNKRISAKNALVHRFF--------RDVTM-PVPPLRL 298


RESULT 13
A41227
protein kinase (EC 2.7.1.37) cdk2 - human
N;Alternate names: Eg1 homolog; protein kinase p34
C;Species: Homo sapiens (man)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C;Accession: A41227; S17873; S16520
R;Ninomiya-Tsuji, J.; Nomoto, S.; Yasuda, H.; Reed, S.I.; Matsumoto, K.
Proc. Natl. Acad. Sci. U.S.A. 88, 9006-9010, 1991
A;Title: Cloning of a human cDNA encoding a CDC2-related kinase by complementation of a
A;Reference number: A41227; MUID:92020980; PMID:1717994
A;Accession: A41227
A;Molecule type: mRNA
A;Residues: 1-298 <NIN>
A;Cross-references: UNIPROT:P24941; GB:M68520; NID:g180177; PIDN:AAA35667.1; PID:g180178
R;Tsai, L.H.; Harlow, E.; Meyerson, M.
Nature 353, 174-177, 1991
A;Title: Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E1A-
A;Reference number: S17873; MUID:91367262; PMID:1653904
A;Accession: S17873
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-298 <TSA>
A;Cross-references: GB:X62071; NID:g312802; PIDN:CAA43985.1; PID:g312803
R;Elledge, S.J.; Spottswood, M.R.
EMBO J. 10, 2653-2659, 1991
A;Title: A new human p34 protein kinase, CDK2, identified by complementation of a cdc28
A;Reference number: S16520; MUID:91330891; PMID:1714386
A;Accession: S16520
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-176,'S',178-298 <ELL>
A;Cross-references: EMBL:X61622; NID:g29848; PIDN:CAA43807.1; PID:g29849
C;Genetics:
A;Gene: GDB:CDK2
A;Cross-references: GDB:128984; OMIM:116953
A;Map position: 12q13-12q13
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine/
F;2-255/Domain: protein kinase homology <KIN>
F;10-18/Region: protein kinase ATP-binding motif
F;14,160/Binding site: phosphate (Thr) (covalent) #status predicted
F;15/Binding site: phosphate (Tyr) (covalent) #status predicted
F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

  Query Match            10.4%;  Score 230;  DB 2;  Length 298;
  Best Local Similarity   28.7%;  Pred. No. 1.9e-06;
  Matches   97;  Conservative   44;  Mismatches  119;  Indels   78;  Gaps   16;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQF-LPPGTTGAAASAAEYGFRKERAALE 81
              :|   ::| |:   ||: |           ||::  |   | |  ::|      :| : |:
Db          4 FQKVEKIGEGTYGVVYKAR--NKLTGEVVALKKIRLDTETEGVPSTAI-----REISLLK 56

Qy         82 QLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARD 141
              :|  | ||| |  |  ||   :    |: | |   : : :  |:  | : :|:
Db         57 EL-NHPNIVKLLDV--IHTENKL---YLVFEFLHQDLKKFMDASALTGIPLPLIKSYLFQ 110

Qy        142 VLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY---IQTDGYR 198
              :|: ||| |    :| ||||:|:| : |    || ||||:   |      |   : |  ||
Db        111 LLQGLAFCHSHRVLHRDLKPQNLLINTEG-AIKLADFGLARAFGVPVRTYTHEVVTLWYR 169

Qy        199 APEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAII 258
              |||  |  |            ::|||:|||| |  ||          |   :  :|   |
Db        170 APEI-LLGC---------KYYSTAVDIWSLGCIFAEM--------VTRRALFPGDSE--I 209

Qy        259 DHIFASKAVVN----------AAIPAYHL--------------------RDLIKSMLHD 287
              | :|    :           ::| |                       | |:  |||
Db        210 DQLFRIFRTLGTPDEVVWPGVTSMPDYKPSFPKWARQDFSKVVPPLDEDGRSLLSQMLHY 269

Qy        288 DPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRL 325
              ||::|| |: ||  |||        :| |  | | |||
Db        270 DPNKRISAKAALAHPFF--------QD-VTKPVPHLRL 298


RESULT 14
S23387
protein kinase (EC 2.7.1.37) cdk6 - human
N;Alternate names: cyclin-dependent kinase 6; protein kinase cdc2-related PLSTIRE
C;Species: Homo sapiens (man)
C;Date: 18-Jun-1993 #sequence_revision 18-Jun-1993 #text_change 09-Jul-2004
C;Accession: S23387; S22749
R;Meyerson, M.; Enders, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.; Tsai
EMBO J. 11, 2909-2917, 1992
A;Title: A family of human cdc2-related protein kinases.
A;Reference number: S23382; MUID:92347325; PMID:1639063
A;Accession: S23387
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-326 <MEY>
A;Cross-references: UNIPROT:Q00534; EMBL:X66365; NID:g36622; PIDN:CAA47008.1; PID:g36623
C;Genetics:
A;Gene: GDB:CDK6
A;Cross-references: GDB:283458
A;Map position: 7q36-7q36
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threonine
F;11-284/Domain: protein kinase homology <KIN>
F;19-27/Region: protein kinase ATP-binding motif
F;24/Binding site: phosphate (Tyr) (covalent) #status predicted
F;43,61,145,147/Active site: Lys, Glu, Asp, Lys #status predicted
F;177/Binding site: phosphate (Thr) (covalent) #status predicted
F;290/Binding site: phosphate (Ser) (covalent) #status predicted

  Query Match            10.4%;  Score 230;  DB 2;  Length 326;
  Best Local Similarity   27.4%;  Pred. No. 2.1e-06;
```

```
  Matches   87;  Conservative   42;  Mismatches  124;  Indels   64;  Gaps   10;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPGTTGAAASAAEYGFRKERAALEQ 82
              ::  : :| |:   |:: |   | |      |  :  | |  |      :| | |
Db         13 YECVAEIGEGAYGKVFKARDLKNGGRFVALKRVRVQTGEEGMPLSTI-----REVAVLRH 67

Qy         83 LQ--GHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCAR 140
              |:   | |:| |: | |:  :       |: | :|  ::  |       |      |:
Db         68 LETFEHPNVVRLFDVCTVSRTDRETKLTLVFEHVDQDLTTYLDKVPEPGVPTETIKDMMF 127

Qy        141 DVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL----SFKEGNQDVKYIQTDG 196
               :|  | |||    || ||||:||| ::  :  || ||||    ||:     |  : |
Db        128 QLLRGLDFLHSHRVVHRDLKPQNILVTSSGQ-IKLADFGLARIYSFQMALTSV--VVTLW 184

Qy        197 YRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSA 256
              |||||  ||:  |          : |||||:| |  |||   |           |:
Db        185 YRAPEVLLQSSYA----------TPVDLWSVGCIFAEMFRRKPL----------FRGSS 223

Qy        257 IIDHIFASKAVV----------NAAIP--AYHL----------------RDLIKSMLHD 287
               :| :     |:          : |:|  |:|                :||:   |
Db        224 DVDQLGKILDVIGLPGEEDWPRDVALPRQAFHSKSAQPIEKFVTDIDELGKDLLLKCLTF 283

Qy        288 DPSRRIPAEMALCSPFF 304
              :|::|| |  ||  |:|
Db        284 NPAKRISAYSALSHPYF 300


RESULT 15
T13024
probable protein kinase (EC 2.7.1.-) F8L21.120 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jan-2000
C;Accession: T13024
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, July 1999
A;Reference number: Z17587
A;Accession: T13024
A;Molecule type: DNA
A;Residues: 1-373 <BEV>
A;Cross-references: EMBL:AL096882; GSPDB:GN00062; ATSP:F8L21.120
A;Experimental source: cultivar Columbia; BAC clone F8L21
C;Genetics:
A;Gene: ATSP:F8L21.120
A;Map position: 4
A;Introns: 58/1; 98/3; 144/3; 255/3; 317/1
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: phosphotransferase; protein kinase
F;40-295/Domain: protein kinase homology <KIN>

  Query Match             10.4%;  Score 230;  DB 2;  Length 373;
  Best Local Similarity   25.9%;  Pred. No. 2.4e-06;
  Matches   83;  Conservative   59;  Mismatches  111;  Indels   68;  Gaps   14;

Qy         75 KERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWM 134
              :|    |  |: | |:| :  :       :     :: ||:|  : :::   |:|  :
Db         86 REIKLLRHLE-HENVVVIKDIIRPPKKEDFVDVYIVFELMDTDLHQII--RSNQSLNDDH 142

Qy        135 IQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL--SFKEGNQDVKYI 192
               |:    :|  | ::|    :| |||| |:| :: |   |: ||||  :  |     :|:
Db        143 CQYFLYQILRGLKYIHSANVLHRDLKPSNLLLNS-NCDLKITDFGLARTTSETEYMTEYV 201

Qy        193 QTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLE------MFSGMKLKHTVR 246
               |  |||||            | : :| |||:|:||:| |  |       :| |    | ::
Db        202 VTRWYRAPEL----------LLNSSEYTSAIDVWSVGCIFAEIMTREPLFPGKDYVHQLK 251

Qy        247 -----------SQEW--KANSSAIIDHI-------FASK--AVVNAAIPAYHLRDLIKS 283
                         | |:   ||:   :  :        |:::  :: : ||      ||::
Db        252 LITELIGSPDGASLEFLRSANARKYVKELPKFPRQNFSARFPSMNSTAI------DLLEK 305

Qy        284 MLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDV 343
              ||  || :||  | ||| |: |      : ||   |         |  : :  : |:
Db        306 MLVFDPVKRITVEEALCYPYLSA-----LHDLNDEP--------VCSNHFSFHFEDPSST 352

Qy        344 VEDVKE----ECQKYGPVVSL 360
               |::||    |  |: |: |:
Db        353 EEEIKELVWLESVKFNPLPSI 373


Search completed: November 29, 2004, 11:42:41
Job time : 27 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 122.5 Seconds
(without alignments)
1968.016 Million cell updates/sec

Title: US-10-798-532-4
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAFG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2195	99.5	419	1	KIST_HUMAN	Q8tas1 homo sapien
2	2183	99.0	419	1	KIST_RAT	Q63285 rattus norv
3	2182	98.9	419	2	AAH58732	Aah58732 mus muscu
4	2182	98.9	419	2	AAO13515	Aao13515 mus muscu
5	2177	98.7	419	1	KIST_MOUSE	P97343 mus musculu
6	1711	77.6	330	2	Q8C0N6	Q8c0n6 mus musculu
7	254	11.5	505	2	Q6CFX3	Q6cfx3 yarrowia li
8	248.5	11.3	882	2	Q8WQL7	Q8wql7 caenorhabdi
9	242	11.0	433	2	Q84SN3	Q84sn3 oryza sativ
10	240.5	10.9	395	2	Q6GM90	Q6gm90 xenopus lae
11	240.5	10.9	720	2	Q9CA22	Q9ca22 arabidopsis
12	240.5	10.9	765	2	Q25758	Q25758 plasmodium
13	240.5	10.9	777	2	Q84VX4	Q84vx4 arabidopsis
14	240.5	10.9	826	2	Q94656	Q94656 plasmodium
15	240.5	10.9	914	2	Q8ILF0	Q8ilf0 plasmodium
16	240	10.9	388	1	SMK1_YEAST	P41808 saccharomyc
17	239.5	10.9	395	2	Q8AYR3	Q8ayr3 xenopus lae
18	239.5	10.9	395	2	AAH70745	Aah70745 xenopus l
19	238.5	10.8	387	1	SKP1_SCHPO	Q10452 schizosacch
20	237.5	10.8	360	2	Q6SZ67	Q6sz67 lytechinus
21	237.5	10.8	360	2	AAR17088	Aar17088 lytechinu
22	235.5	10.7	434	2	Q75B85	Q75b85 ashbya goss
23	235.5	10.7	434	2	AAS51605	Aas51605 ashbya go
24	235	10.7	298	1	CDK2_MESAU	P48963 mesocricetu
25	234.5	10.6	392	2	Q6ZD93	Q6zd93 oryza sativ
26	234.5	10.6	392	2	BAC99508	Bac99508 oryza sat
27	234.5	10.6	433	1	KKQ1_YEAST	P36005 saccharomyc
28	234	10.6	376	1	MPK4_ARATH	Q39024 arabidopsis
29	234	10.6	376	2	AAM66070	Aam66070 arabidops
30	233.5	10.6	433	2	Q6BG29	Q6bg29 paramecium
31	233.5	10.6	1233	2	Q6CAC2	Q6cac2 yarrowia li
32	233	10.6	369	2	Q7M445	Q7m445 dictyosteli
33	232.5	10.5	367	1	MK12_RAT	Q63538 rattus norv
34	232	10.5	568	2	Q9HBY6	Q9hby6 homo sapien
35	232	10.5	588	2	Q9HBY7	Q9hby7 homo sapien
36	232	10.5	588	2	AAT06103	Aat06103 homo sapi
37	231.5	10.5	367	1	MK12_MOUSE	O08911 mus musculu
38	231.5	10.5	464	1	MK10_RAT	P49187 rattus norv
39	231.5	10.5	542	2	Q6I587	Q6i587 oryza sativ
40	231	10.5	298	1	CDK2_CRIGR	O55076 cricetulus
41	231	10.5	298	1	CDK2_RAT	Q63699 rattus norv
42	231	10.5	298	2	Q6P751	Q6p751 rattus norv
43	231	10.5	298	2	AAH61832	Aah61832 rattus no
44	231	10.5	361	1	JNK_SUBDO	Q966y3 suberites d
45	231	10.5	435	1	HOG1_YEAST	P32485 saccharomyc

ALIGNMENTS

RESULT 1
KIST_HUMAN
ID KIST_HUMAN STANDARD; PRT; 419 AA.
AC Q8TAS1; Q96C22;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting
DE with stathmin).
GN Name=KIST; Synonyms=KIS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX PubMed=12782393;
RA Bieche I., Manceau V., Curmi P.A., Laurendeau I., Lachkar S.,
RA Leroy K., Vidaud D., Sobel A., Maucuer A.;
RT "Quantitative RT-PCR reveals a ubiquitous but preferentially neural
RT expression of the KIS gene in rat and human.";
RL Brain Res. Mol. Brain Res. 114:55-64(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Eye, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Possesses kinase activity. May be involved in
CC trafficking and/or processing of RNA (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Interacts with stathmin (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;

```
CC         IsoId=Q8TAS1-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q8TAS1-2; Sequence=VSP_004908, VSP_004909;
CC         Note=No experimental confirmation available;
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC   -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ536197; CAD60192.1; -.
DR   EMBL; BC014917; AAH14917.1; -.
DR   EMBL; BC026046; AAH26046.1; -.
DR   HSSP; P26368; 1O0P.
DR   GO; GO:0005737; C:cytoplasm; ISS.
DR   GO; GO:0005634; C:nucleus; ISS.
DR   GO; GO:0005524; F:ATP binding; ISS.
DR   GO; GO:0005515; F:protein binding; ISS.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
DR   GO; GO:0003723; F:RNA binding; ISS.
DR   GO; GO:0018105; P:peptidyl-serine phosphorylation; ISS.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   Pfam; PF00076; RRM_1; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00360; RRM; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR   PROSITE; PS50102; RRM; 1.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW   Alternative splicing; ATP-binding; Nuclear protein; RNA-binding;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       23    304       Protein kinase.
FT   DOMAIN      324    406       RNA-binding (RRM).
FT   NP_BIND      29     37       ATP (By similarity).
FT   BINDING      54     54       ATP (By similarity).
FT   ACT_SITE    158    158       Proton acceptor (By similarity).
FT   VARSPLIC    342    344       DVV -> GLC (in isoform 2).
FT                                /FTId=VSP_004908.
FT   VARSPLIC    345    419       Missing (in isoform 2).
FT                                /FTId=VSP_004909.
SQ   SEQUENCE   419 AA;  46546 MW;  903E982EE12A8CF8 CRC64;

  Query Match             99.5%;  Score 2195;  DB 1;  Length 419;
  Best Local Similarity   99.5%;  Pred. No. 2.4e-158;
  Matches  417;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEEEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 2
KIST_RAT
ID   KIST_RAT       STANDARD;      PRT;   419 AA.
AC   Q63285;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting
DE   with stathmin) (PAM COOH-terminal interactor protein 2) (P-CIP2).
GN   Name=Kist; Synonyms=Kis;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, TISSUE
RP   SPECIFICITY, AND MUTAGENESIS OF LYS-54.
RC   TISSUE=Striatum;
RX   MEDLINE=97435279; PubMed=9287318;
RA   Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P.,
RA   Sobel A.;
RT   "KIS is a protein kinase with an RNA recognition motif.";
RL   J. Biol. Chem. 272:23151-23156(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Caldwell B.D., Darlington D.N., Penzes P., Johnson R.C., Eipper B.A.,
RA   Mains R.E.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 28-419 FROM N.A.
RC   TISSUE=Hippocampus;
RX   MEDLINE=97067094; PubMed=8910496;
RA   Alam R., Caldwell B.D., Johnson R.C., Darlington D.N., Mains R.E.,
RA   Eipper B.A.;
RT   "Novel proteins that interact with the COOH-terminal cytosolic routing
RT   determinants of an integral membrane peptide-processing enzyme.";
RL   J. Biol. Chem. 271:28636-28640(1996).
CC   -!- FUNCTION: Possesses kinase activity. May be involved in
CC       trafficking and/or processing of RNA.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBUNIT: Interacts with stathmin (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC   -!- TISSUE SPECIFICITY: In the embryo, preferentially expressed in the
CC       developing nervous system.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC   -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X98374; CAA67021.1; -.
DR   EMBL; U70372; AAC53031.2; -.
DR   HSSP; P26368; 1O0P.
DR   RGD; 2968; Kist.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
```

```
DR   GO; GO:0005524; F:ATP binding; IC.
DR   GO; GO:0005515; F:protein binding; ISS.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
DR   GO; GO:0003723; F:RNA binding; NAS.
DR   GO; GO:0018105; P:peptidyl-serine phosphorylation; IDA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   Pfam; PF00076; RRM_1; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00360; RRM; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR   PROSITE; PS50102; RRM; 1.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW   ATP-binding; Nuclear protein; RNA-binding;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       23    304       Protein kinase.
FT   DOMAIN      324    406       RNA-binding (RRM).
FT   NP_BIND      29     37       ATP (By similarity).
FT   BINDING      54     54       ATP.
FT   ACT_SITE    141    141       Proton acceptor (By similarity).
FT   ACT_SITE    158    158       Proton acceptor (By similarity).
FT   MUTAGEN      54     54       K->R: Loss of activity.
SQ   SEQUENCE   419 AA;  46505 MW;  D2F9B7BF8F080DF3 CRC64;

  Query Match             99.0%;  Score 2183;  DB 1;  Length 419;
  Best Local Similarity   98.8%;  Pred. No. 1.9e-157;
  Matches  414;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db          1 MAGSGCAWGAEPPRFLEAPGRLWQVQSRLGSGSSASVYRVRCCGTPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||||| ||:||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 3
AAH58732
ID   AAH58732    PRELIMINARY;      PRT;   419 AA.
AC   AAH58732;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Kinase interacting with leukemia-associated gene (Stathmin).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC058732; AAH58732.1; -.
KW   Kinase.
SQ   SEQUENCE   419 AA;  46489 MW;  3BD3C06A59E22540 CRC64;

  Query Match             98.9%;  Score 2182;  DB 2;  Length 419;
  Best Local Similarity   98.8%;  Pred. No. 2.3e-157;
  Matches  414;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGTPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||||| ||:||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 4
AAO13515
ID   AAO13515    PRELIMINARY;      PRT;   419 AA.
AC   AAO13515;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
```

```
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   KIS kinase.
GN   KIS.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129S6;
RA   Crook M.F., Boehm M., Nabel E.G.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY180177; AAO13515.1; -.
KW   Kinase.
SQ   SEQUENCE   419 AA;  46489 MW;  3BD3C06A59E22540 CRC64;

  Query Match             98.9%;  Score 2182;  DB 2;  Length 419;
  Best Local Similarity   98.8%;  Pred. No. 2.3e-157;
  Matches  414;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||:||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGTPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||||| ||:|||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 5
KIST_MOUSE
ID   KIST_MOUSE     STANDARD;      PRT;   419 AA.
AC   P97343; Q61775; Q9CYT1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting
DE   with stathmin).
GN   Name=Kist; Synonyms=KIS;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [2]
RP   SEQUENCE OF 1-414 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=97435279; PubMed=9287318;
RA   Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P.,
RA   Sobel A.;
RT   "KIS is a protein kinase with an RNA recognition motif.";
RL   J. Biol. Chem. 272:23151-23156(1997).
RN   [3]
RP   REVISIONS TO 52 AND 68-69.
RA   Maucuer A.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 122-321 FROM N.A., AND INTERACTION WITH STATHMIN.
RC   TISSUE=Embryo;
RX   MEDLINE=95241452; PubMed=7724523;
RA   Maucuer A., Camonis J.H., Sobel A.;
RT   "Stathmin interaction with a putative kinase and coiled-coil-forming
RT   protein domains.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:3100-3104(1995).
CC   -!- FUNCTION: Possesses kinase activity. May be involved in
CC       trafficking and/or processing of RNA (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBUNIT: Interacts with stathmin.
CC   -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC   -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Y10725; CAA71714.2; -.
DR   EMBL; AK013347; BAB28802.1; -.
DR   EMBL; X82320; CAA57763.1; -.
DR   PIR; I48615; I48615.
DR   HSSP; P26368; 1OOP.
DR   MGD; MGI:1341908; Kist.
DR   GO; GO:0005737; C:cytoplasm; ISS.
DR   GO; GO:0005634; C:nucleus; ISS.
DR   GO; GO:0005524; F:ATP binding; ISS.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
DR   GO; GO:0003723; F:RNA binding; ISS.
DR   GO; GO:0018105; P:peptidyl-serine phosphorylation; ISS.
DR   InterPro; IPR011009; Kinase_like.
```

```
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   Pfam; PF00076; RRM_1; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00360; RRM; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR   PROSITE; PS50102; RRM; 1.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW   ATP-binding; Nuclear protein; RNA-binding;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       23    303       Protein kinase.
FT   DOMAIN      323    405       RNA-binding (RRM).
FT   NP_BIND      29     37       ATP (By similarity).
FT   BINDING      54     54       ATP (By similarity).
FT   ACT_SITE    158    158       Proton acceptor (By similarity).
FT   CONFLICT    170    170       N -> D (in Ref. 4).
FT   CONFLICT    231    231       T -> I (in Ref. 3 and 4).
SQ   SEQUENCE   419 AA;  46477 MW;  A91B307BA50F29A1 CRC64;

  Query Match             98.7%;  Score 2177;  DB 1;  Length 419;
  Best Local Similarity   98.6%;  Pred. No. 5.5e-157;
  Matches  413;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||:||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGTPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGITLLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPGRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||||| ||:||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 6
Q8C0N6
ID   Q8C0N6      PRELIMINARY;      PRT;   330 AA.
AC   Q8C0N6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE   library, clone:4933401D07 product:kinase interacting with leukemia-
DE   associated gene (stathmin), full insert sequence.
GN   Name=Kist;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=99279253; PubMed=10349636;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=20499374; PubMed=11042159;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=20530913; PubMed=11076861;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK030152; BAC26812.1; -.
DR   HSSP; P26368; 1O0P.
DR   MGD; MGI:1341908; Kist.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004672; F:protein kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00069; Pkinase; 2.
DR   Pfam; PF00076; RRM_1; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00360; RRM; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
```

```
DR   PROSITE; PS50102; RRM; 1.
KW   Kinase.
SQ   SEQUENCE   330 AA;  37201 MW;  4164C0758EE7CB4C CRC64;

  Query Match             77.6%;  Score 1711;  DB 2;  Length 330;
  Best Local Similarity   98.8%;  Pred. No. 1.1e-121;
  Matches  326;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         90 VTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFL 149
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFL 60

Qy        150 HHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLA 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLA 120

Qy        210 QAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVN 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVN 180

Qy        270 AAIPAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVL 329
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 AAIPAYHLRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVL 240

Qy        330 DDDYLGNEEEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLL 389
              ||||| ||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DDDYLENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLL 300

Qy        390 TGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||
Db        301 TGRMFDGKFVVATFYPLSAYKRGYLYQTLL 330


RESULT 7
Q6CFX3
ID   Q6CFX3      PRELIMINARY;      PRT;   505 AA.
AC   Q6CFX3;
DT   01-OCT-2004 (TrEMBLrel. 28, Created)
DT   01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Chromosome B of strain CLIB99 of Yarrowia lipolytica.
GN   ORFNames=YALI0B02816g;
OS   Yarrowia lipolytica (Candida lipolytica).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Dipodascaceae; Yarrowia.
OX   NCBI_TaxID=4952;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIB99;
RG   GENOLEVURES;
RA   Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA   Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA   Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA   Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA   Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA   Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA   Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA   Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA   Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA   Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA   Swennene D., Tekaia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA   Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA   Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA   Wincker P., Souciet J.L.;
RT   "Genome evolution in yeasts.";
RL   Nature 430:35-44(2004).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIB99;
RA   Genoscope;
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR382128; CAG82657.1; -.
SQ   SEQUENCE   505 AA;  58191 MW;  D30933013FCEE357 CRC64;

  Query Match             11.5%;  Score 254;  DB 2;  Length 505;
  Best Local Similarity   25.2%;  Pred. No. 9.5e-11;
  Matches   87;  Conservative   47;  Mismatches  129;  Indels   82;  Gaps   11;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPGTTGAAASAAEYGFRKERAALEQ 82
              :||   || |:   |   :  | |  |  |:::    :       |    :|   |
Db         22 FQVTKELGHGAYGIVCAAKYTGTPDGPGVAIKKVTNIFSKNILCKRA----LRELKLLNH 77

Qy         83 LQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDV-SVSELLLYSSHQGCSMWMI------ 135
               :||:||  ||           ::::|  : :|| ||     | |  |
Db         78 FRGHKNITCLYD----------------MDIVDTNNFNELYLYEELMECDMHQIIRSGQP 121

Qy        136 ------QHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDV 189
                   |     :|  : ::|    :| |||| |:| :|: |  |: ||||:    : |
Db        122 LTDAHYQSFVYQILAGVKYIHSADVLHRDLKPGNLLVNADCE-LKICDFGLARGFSSDDE 180

Qy        190 K-------YIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLK 242
              |       |: |  ||||| |           |     | |:|:||:| || |:  |   |
Db        181 KNAGFLTEYVATRWYRAPEIML----------SFQSYTKAIDIWSVGCILAELLGG---K 227

Qy        243 HTVRSQEWKANSSAIIDHI----------------------FASKAVVNAAIPAYH-- 276
               : : :   : |: ::                        |  |   :   |  :
Db        228 PLFKGKNYVDQLNQILHYLGTPSEETLKRIGSPRAQEYVRGLPFMPKIPFSTLFPTANPE 287

Qy        277 LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTP 321
               ||:: ||  ||: |:  | ||  |:  |   |   |  : |||
Db        288 ALDLLEKMLAFDPAERVTVEEALEHPYLKIWHDP--RDEPVCPTP 330


RESULT 8
Q8WQL7
ID   Q8WQL7      PRELIMINARY;      PRT;   882 AA.
AC   Q8WQL7; Q22155;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Minibrain kinase (Hypothetical protein T04C10.1).
GN   Name=mbk-1; Synonyms=T04C10.1;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22505392; PubMed=12618396;
RA   Raich W.B., Moorman C., Lacefield C.O., Lehrer J., Bartsch D.,
RA   Plasterk R.H., Kandel E.R., Hobert O.;
RT   "Characterization of Caenorhabditis elegans homologs of the Down
RT   syndrome candidate gene DYRK1A.";
RL   Genetics 163:571-580(2003).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Burton J.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY064464; AAL40874.1; -.
DR   EMBL; Z69885; CAA93756.2; -.
DR   HSSP; P24941; 1H0V.
DR   IntAct; Q8WQL7; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
```

```
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Hypothetical protein; Kinase;
KW   Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   882 AA;  101244 MW;  DD5C9DD354DA7098 CRC64;

  Query Match             11.3%;  Score 248.5;  DB 2;  Length 882;
  Best Local Similarity   25.7%;  Pred. No. 5e-10;
  Matches  104;  Conservative   57;  Mismatches  113;  Indels  131;  Gaps   17;

Qy         88 NIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALA 147
              ||||| | | :| :      ||: |||  :: :||  :| :| |: : :  |: : : |
Db        389 NIVTLKGHF-VHRA----HLCLVFELLSYNLYDLLKNTSFRGVSLNLARKFAQQLGKTLL 443

Qy        148 FLHHE--GYVHADLKPRNILW-SAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAEL 204
              ||       :| |||| |:|  :|:    ::|||| | : |::  :|||:  ||:||  |
Db        444 FLSSPELSIIHCDLKPENVLLVNAKRSQIRVIDFGSSCQTGHRIYQYIQSRFYRSPEVLL 503

Qy        205 QNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSG-------------MKL---------- 241
                    |:  ||:    :|:|||| ||:|| :|             ||:
Db        504 -------GIAYDTK----IDMWSLGCILVEMHTGEPLFAGSSEVDQMMKIVEVLGMPPKE 552

Qy        242 ----------------------------KHTVRSQEWKANSSAIIDHIFASKAVVNAAI 272
                                         :||     :||  :  :  |       | :
Db        553 MLDIGPKTHKYFDKTEDGIYYCKKTRDGYRHT-----YKAPGARKLHEILG----VTSGG 603

Qy        273 PA--------------YHLRDLIKSMLHDDPSRRIPAEMALCSPFFSI--PFAPHIEDLV 316
              |                    :|||| ||  || :||        ||:: :  ||    |: |
Db        604 PGGRRLGEPGHSVEDYSKFKDLIKRMLQFDPKQRI-------SPYYVVRHPFLKQKEERV 656

Qy        317 MLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKY--GPVVSLLVPKENPGRGQVFV 374
                ||                          ::::: | |   |   :    |:|  | |:|
Db        657 PSQPPVSH--------------------SNLQQQQQLYIQQPSSQMSQVMESPSVGSVYV 696

Qy        375 E----YANAGDSKAAQKLLTGRMFDGKF--VVATFYPLSAYKRGY 413
              |    |  | | |    :|    :|    | |     ||::: |
Db        697 EDNGMYRQAPGSSANPISVTSSFDEGDAMEVDAGRRRFSAHQQNY 741


RESULT 9
Q84SN3
ID   Q84SN3      PRELIMINARY;      PRT;   433 AA.
AC   Q84SN3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein OSJNBb0043P23.8.
GN   Name=OSJNBb0043P23.8;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA   Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA   Fadrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA   Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA   Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA   White O., Salzberg S.L., Fraser C.M.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Buell R.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AC099324; AAO73290.1; -.
DR   HSSP; P24941; 1E1X.
DR   Gramene; Q84SN3; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Hypothetical protein; Kinase;
KW   Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   433 AA;  49515 MW;  28AC98AE930BDE89 CRC64;

  Query Match             11.0%;  Score 242;  DB 2;  Length 433;
  Best Local Similarity   26.5%;  Pred. No. 6.4e-10;
  Matches   99;  Conservative   54;  Mismatches  138;  Indels   82;  Gaps   15;

Qy         75 KERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWM 134
              :|  ||::|  | ||| |  |   :          :   | || : |     | :    :|
Db         49 REVKALQKL-NHPNIVKLKEVTMENHELFFIFENMECNLYDV-IRERQAAFSEEEIRNFM 106

Qy        135 IQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQD--VKYI 192
              :|     :|: ||::|: || | |||| |:|    :   |: ||||: :  :      |:
Db        107 VQ-----ILQGLAYMHNNGYFHRDLKPENLL--VTDGTVKIADFGLAREVSSSPPYTDYV 159

Qy        193 QTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL---------- 241
              |  |||||  ||:          :  | |:|:|::| || |:|:    |
Db        160 STRWYRAPEVLLQS----------SAYTPAIDMWAVGAILAELFTLSPLFPGGSETDQLY 209

Qy        242 ---------KHTVRSQEWKANSSAIIDHIFAS-----KAVVNAAIPAYHLRDLIKSMLHD 287
                        |||  :      |:  :          : : || : |    |||: :
Db        210 KICAVLGTPDHTVWPEGMNLPRSSSFNFFQIPPRNLWELIPNATLEAI---DLIQQLCSW 266

Qy        288 DPSRRIPAEMALCSPFFSIP-------FAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEE- 339
              || ||  || :|  |||::          | | : :   | | |  ||: |    | | |
Db        267 DPRRRPTAEQSLQHPFFNVGNWVPRPLHASHTKTIETRPNPRLE-LNLWD---FGTEPED 322

Qy        340 -YEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKF 398
               | |:   :|                ||      :::|       :::|     |:
Db        323 NYLDLTLSLKPSF---------------PG-----TDFSNNVPEHTKEEILLYPGFENPP 362

Qy        399 VVATFYPLSAYKR 411
              | : |:|| |  |
Db        363 VQSGFWPLVASDR 375


RESULT 10
Q6GM90
ID   Q6GM90      PRELIMINARY;      PRT;   395 AA.
AC   Q6GM90;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
```

```
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22341132; PubMed=12454917;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Klein S., Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC074183; AAH74183.1; -.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Hypothetical protein; Kinase;
KW   Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   395 AA;  43896 MW;  4C78B8BF828A2CC4 CRC64;

  Query Match             10.9%;  Score 240.5;  DB 2;  Length 395;
  Best Local Similarity   27.4%;  Pred. No. 7.5e-10;
  Matches   90;  Conservative   48;  Mismatches  108;  Indels   83;  Gaps   16;

Qy          4 SGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASV-----------YRVRCCGNPGSPPGA 52
              ||  |  :    :|   |::| :  ||:|: : |           : |:|      |  |
Db          9 SGACWKKQ----VEDIKRMFQFKEVLGTGAFSEVVLAEEKETGKLFAVKCI-----PKKA 59

Qy         53 LRQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLL- 111
              |:                    |     | | | ::: | ||| |  ::          ||   |:
Db         60 LK------------GKESSIENEIAVLRKIK-HENIVALEDIY------ESPSHLYLVM 99

Qy        112 ------ELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNIL 165
                   || |  : |   |:     ::       | ||:|:::||  | || |||| |:|
Db        100 QLVSGGELFD-RIVEKGFYTEKDASTL------IRQVLDAVSYLHRLGIVHRDLKPENLL 152

Qy        166 WSAENECFKLI--DFGLSFKEGNQDVKYIQ--TDGYRAPEAELQNCLAQAGLQSDTECTS 221
              : :: |  |::  |||||  ||  ||      | || |||      |||         :
Db        153 YFSQVEESKIMISDFGLSKMEGKGDVMSTACGTPGYVAPEV-----LAQ------KPYSK 201

Qy        222 AVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAY-----H 276
              ||| ||:|:|   :  |           :  | | : : |   ||      | :

Db        202 AVDCWSIGVIAYILLCGY-------PPFYDENDSRLFEQIL--KADYEFDSPYWDDISES 252

Qy        277 LRDLIKSMLHDDPSRRIPAEMALCSPFFS 305
                :| |::::  ||::|   | ||  |: :
Db        253 AKDFIQNLMEKDPNKRYTCEQALRHPWIA 281


RESULT 11
Q9CA22
ID   Q9CA22      PRELIMINARY;      PRT;   720 AA.
AC   Q9CA22;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein T32E8.5.
GN   Name=T32E8.5;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA   Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA   Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Town C.D., Kaul S.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC012193; AAG51619.1; -.
DR   PIR; A96807; A96807.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW   ATP-binding; Hypothetical protein; Kinase; Transferase.
SQ   SEQUENCE   720 AA;  80027 MW;  F26A9A0B1C6D91DE CRC64;

  Query Match             10.9%;  Score 240.5;  DB 2;  Length 720;
  Best Local Similarity   25.7%;  Pred. No. 1.6e-09;
  Matches  105;  Conservative   57;  Mismatches  146;  Indels  101;  Gaps   17;

Qy         11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPGTTGAAASAAE 70
              :|  | :  |:|:|   ::||| |: |::|    :      ||::    |   | ||
Db        331 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 383

Qy         71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
              ||| :|   |::|:|  ||: |           |  : || |:|: ::|
Db        384 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 436

Qy        121 ---LLLYS----SHQGCSMW------------MIQHCARDVLEALAFLHHEGYVHADLK 160
                 :| |     :|      |                ::   : :|:|:  :| |  ||:|||
Db        437 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 496

Qy        161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
              | | |          ||||||:: |  | |   || |         | :||| : |   : |
Db        497 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 551

Qy        215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272
              :  :|    |:|||| || :|  | :          | |     | |  |       : |
Db        552 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 607
```

```
Qy        273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH---------------------- 311
              | ||:|  |  | ::|      |  || : |  ||
Db        608 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 663

Qy        312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351
                      :| |:  | |: |      :| | : :: : ::  | | |
Db        664 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 706


RESULT 12
Q25758
ID   Q25758      PRELIMINARY;      PRT;   765 AA.
AC   Q25758;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mitogen-activated protein kinase-related protein.
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Dd2;
RA   Lin D.T., Goldman N.D., Syin C.;
RT   "Stage-specific expression of a Plasmodium falciparum protein related
RT   to the eukaryotic mitogen-activated protein kinases.";
RL   Mol. Biochem. Parasitol. 78:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Dd2;
RA   Lin D.T., Goldman N.D., Syin C.;
RL   Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; U36377; AAC47170.1; -.
DR   HSSP; Q16539; 1KV1.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR003527; MAP_kin.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS01351; MAPK; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   765 AA;  90231 MW;  5BEC8EA6A4B02D32 CRC64;

  Query Match             10.9%;  Score 240.5;  DB 2;  Length 765;
  Best Local Similarity   25.1%;  Pred. No. 1.7e-09;
  Matches  103;  Conservative   57;  Mismatches  145;  Indels  105;  Gaps   19;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPP-----GALRQFLPPGTTGAAASAAEYGFRKER 77
              : :  ::| |:   |:: ||  |          || :           : |:  || | |
Db         23 YDILKKVGKGAYGVVFKGRCKKNKNIVAVKKIFGAFQN----------CTDAQRTFR-EI 71

Qy         78 AALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSEL----LLYSSHQGCSMW 133
                | :| || ||: |  |      : |      |: : ::  : |:    ||   |:
Db         72 IFLYELNGHDNIIKLMDVIK---AKNDNDIYLIFDFMETDLHEVIKADLLEEIHK----- 123

Qy        134 MIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEGN 186
               ::    :| || ::|  | :| |:|| |||    :||  |: ||||:          |
Db        124 --KYIIYQLLRALKYIHSGGLLHRDIKPSNIL--VNSECHIKVADFGLARSISTHVNENK 179

Qy        187 QDV--KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL--- 241
                  :   |: |  |||||            |   | |  ||:|||| |: |:  |  |
```

```
Db        180 VPILTDYVATRWYRAPEI----------LLGSTHYTEDVDMWSLGCIMGELLCGKPLFTG 229

Qy        242 KHTVRSQE------WKANSSAIID--HIFASKAVVN-AAIPAYHLR----------DLI 281
              |:  |       | |  | |    || | : :   :  :|:           ||:
Db        230 NSTMNQLEKIIQVIGKPNKKDIEDIRSPFAEKIISSFVDLKKKNLKDICYKASNESLDLL 289

Qy        282 KSMLHDDPSRRIPAEMAL----CSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNE 337
              : :|  :||:|| || ||       | ||   |    :: :|
Db        290 EKLLQFNPSKRISAENALKHKYVEEFHSIIDEPTCRHIITIP------------------ 331

Qy        338 EEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQK 387
                 : :: |     |  ||  :: : |      |     | |:||  :|
Db        332 -----INDNTKYRVNFYRNVVYFVIMRRNKFHSNVL----NQGESKKEEK 372


RESULT 13
Q84VX4
ID   Q84VX4      PRELIMINARY;      PRT;   777 AA.
AC   Q84VX4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   At1g77720.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA   Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA   Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA   Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA   Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA   Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA   Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT004636; AAO42882.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW   ATP-binding; Kinase; Transferase.
SQ   SEQUENCE   777 AA;  86376 MW;  F2C25DA9609A878F CRC64;

  Query Match             10.9%;  Score 240.5;  DB 2;  Length 777;
  Best Local Similarity   25.7%;  Pred. No. 1.7e-09;
  Matches  105;  Conservative   57;  Mismatches  146;  Indels  101;  Gaps   17;

Qy         11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPGTTGAAASAAE 70
              :|  | :  |:|:|    ::||| |: |::|     :       ||::     |   | |
Db        388 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 440

Qy         71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
              ||| :|   |::|:|  ||: |              |  : || |:|: ::|
Db        441 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 493

Qy        121 ---LLLYS----SHQGCSMW------------MIQHCARDVLEALAFLHHEGYVHADLK 160
                 :| |     :|     |                ::   : :|:|:  :| |  ||:|||
Db        494 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 553

Qy        161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
```

```
Db        554 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 608

Qy        215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272

Db        609 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 664

Qy        273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH---------------------- 311

Db        665 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 720

Qy        312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351

Db        721 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 763


RESULT 14
Q94656
ID   Q94656      PRELIMINARY;      PRT;   826 AA.
AC   Q94656;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mitogen-activated protein kinase 1, serine/threonine protein
DE   kinase.
GN   Name=pfmap1;
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IT04;
RX   MEDLINE=97080491; PubMed=8921836;
RA   Doerig C.D., Parzy D., Langsley G., Horrocks P., Carter R.,
RA   Doerig C.D.;
RT   "A map kinase homologue from the human malaria parasite Plasmodium
RT   falciparum.";
RL   Gene 177:1-6(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IT04;
RA   Doerig C.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; X82646; CAA57972.1; -.
DR   PIR; JC5153; JC5153.
DR   HSSP; Q16539; 1KV1.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR003527; MAP_kin.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS01351; MAPK; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   826 AA;  96645 MW;  A668EF77E734A38A CRC64;

  Query Match           10.9%;  Score 240.5;  DB 2;  Length 826;
  Best Local Similarity   25.1%;  Pred. No. 1.9e-09;
  Matches  103;  Conservative   57;  Mismatches  145;  Indels  105;  Gaps   19;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPP-----GALRQFLPPGTTGAAASAAEYGFRKER 77

Db         23 YDILKKVGKGAYGVVFKGRCKKNKNIVAVKKIFGAFQN----------CTDAQRTFR-EI 71

Qy         78 AALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSEL----LLYSSHQGCSMW 133

Db         72 IFLYELNGHDNIIKLMDVIK---AKNDNDIYLIFDFMETDLHEVIKADLLEEIHK----- 123

Qy        134 MIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEGN 186

Db        124 --KYIIYQLLRALKYIHSGGLLHRDIKPSNIL--VNSECHIKVADFGLARSISTHVNENK 179

Qy        187 QDV--KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL--- 241

Db        180 VPILTDYVATRWYRAPEI----------LLGSTHYTEDVDMWSLGCIMGELLCGKPLFTG 229

Qy        242 KHTVRSQE------WKANSSAIID--HIFASKAVVN-AAIPAYHLR-----------DLI 281

Db        230 NSTMNQLEKIIQVIGKPNKKDIEDIRSPFAEKIISSFVDLKKKNLKDICYKASNESLDLL 289

Qy        282 KSMLHDDPSRRIPAEMAL----CSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNE 337

Db        290 EKLLQFNPSKRISAENALKHKYVEEFHSIIDEPTCRHIITIP----------------- 331

Qy        338 EEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQK 387

Db        332 -----INDNTKYRVNFYRNVVYFVIMRRNKFHSNVL----NQGESKKEEK 372


RESULT 15
Q8ILF0
ID   Q8ILF0      PRELIMINARY;      PRT;   914 AA.
AC   Q8ILF0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mitogen-activated protein kinase 1.
GN   ORFNames=PF14_0294;
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22255705; PubMed=12368864;
RA   Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA   Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA   Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA   Chan M.S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA   Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA   Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA   McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA   Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA   Fraser C.M., Barrell B.;
RT   "Genome sequence of the human malaria parasite Plasmodium
RT   falciparum.";
RL   Nature 419:498-511(2002).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AE014820; AAN36907.1; -.
DR   HSSP; Q16539; 1KV1.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR003527; MAP_kin.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS01351; MAPK; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   914 AA;  107277 MW;  CA4F486AB051F472 CRC64;
```

```
  Query Match           10.9%;  Score 240.5;  DB 2;  Length 914;
  Best Local Similarity  25.1%;  Pred. No. 2.1e-09;
  Matches  103;  Conservative   57;  Mismatches  145;  Indels  105;  Gaps   19;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPP-----GALRQFLPPGTTGAAASAAEYGFRKER 77
              : :  ::| |:   |:: ||  |          || :          : |:  || |
Db         23 YDILKKVGKGAYGVVFKGRCKKNKNIVAVKKIFGAFQN----------CTDAQRTFR-EI 71

Qy         78 AALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSEL----LLYSSHQGCSMW 133
               | :| || ||: |  |     : |     |: : ::  : |:    ||   |:
Db         72 IFLYELNGHDNIIKLMDVIK---AKNDNDIYLIFDFMETDLHEVIKADLLEEIHK----- 123

Qy        134 MIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEGN 186
                ::    :| || ::|  | :| |:|| |||    :||  |: ||||:        |
Db        124 --KYIIYQLLRALKYIHSGGLLHRDIKPSNIL--VNSECHIKVADFGLARSISTHVNENK 179

Qy        187 QDV--KYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKL--- 241
              :   |: |  |||||           |   |  |  ||:|||| |: |:  |  |
Db        180 VPILTDYVATRWYRAPEI----------LLGSTHYTEDVDMWSLGCIMGELLCGKPLFTG 229

Qy        242 KHTVRSQE------WKANSSAIID--HIFASKAVVN-AAIPAYHLR----------DLI 281
              |:   |       | |   | |    || | : :   :   :|:          ||:
Db        230 NSTMNQLEKIIQVIGKPNKKDIEDIRSPFAEKIISSFVDLKKKNLKDICYKASNESLDLL 289

Qy        282 KSMLHDDPSRRIPAEMAL----CSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNE 337
              : :|  :||:|| || ||       | ||   |    :: :|
Db        290 EKLLQFNPSKRISAENALKHKYVEEFHSIIDEPTCRHIITIP----------------- 331

Qy        338 EEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQK 387
                : :: |      |  ||  :: : |     |     | |:||  :|
Db        332 -----INDNTKYRVNFYRNVVYFVIMRRNKFHSNVL----NQGESKKEEK 372

Search completed: November 29, 2004, 11:46:53
Job time : 123.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 109.5 Seconds
(without alignments)
1372.672 Million cell updates/sec

Title: US-10-798-532-4
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAPG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2195	99.5	419	4	AAG65767	Aag65767 Human pro
2	2195	99.5	419	5	ABP64797	Abp64797 Human pro
3	2190.5	99.3	418	3	AAY70306	Aay70306 Human tra
4	2188	99.2	419	4	AAU03534	Aau03534 Human pro
5	2188	99.2	419	4	AAE11769	Aae11769 Human kin
6	2187.5	99.2	418	3	AAY70305	Aay70305 Human wil
7	2183	99.0	419	7	ADD48093	Add48093 Rat Prote
8	483.5	21.9	137	5	ABP64662	Abp64662 Human ORF
9	240.5	10.9	638	3	AAG38502	Aag38502 Arabidops
10	240.5	10.9	720	3	AAG38501	Aag38501 Arabidops
11	240.5	10.9	777	3	AAG38500	Aag38500 Arabidops
12	236.5	10.7	544	2	AAW85028	Aaw85028 CDK2-gree
13	235	10.7	298	7	ADD21398	Add21398 Golden ha
14	234.5	10.6	433	3	AAB19094	Aab19094 Protein e
15	234.5	10.6	433	8	ADP87472	Adp87472 S cerevis
16	234	10.6	376	3	AAG04672	Aag04672 Arabidops
17	233	10.6	435	6	AAE31916	Aae31916 Yeast Hog
18	232.5	10.5	367	3	AAB22137	Aab22137 Rat stres
19	232	10.5	435	6	AAE31917	Aae31917 Yeast Hog
20	232	10.5	435	6	AAE31922	Aae31922 Yeast Hog
21	232	10.5	435	6	AAE31921	Aae31921 Yeast Hog
22	232	10.5	435	6	AAE31914	Aae31914 Yeast Hog
23	232	10.5	435	6	AAE31913	Aae31913 Yeast Hog
24	232	10.5	544	2	AAW85029	Aaw85029 Green flu
25	232	10.5	568	2	AAW75791	Aaw75791 Human pro
26	232	10.5	568	3	AAB00059	Aab00059 hYAK3-2 p
27	232	10.5	568	4	AAE02013	Aae02013 Human YAK
28	232	10.5	587	7	ABR57639	Abr57639 Human DYR
29	232	10.5	588	2	AAW75790	Aaw75790 Human pro
30	232	10.5	588	3	AAY68792	Aay68792 Amino aci
31	232	10.5	588	3	AAB00058	Aab00058 hYAK3-2 a
32	232	10.5	588	4	AAE02012	Aae02012 Human YAK
33	232	10.5	801	7	ADD93285	Add93285 p53-Yak3
34	231.5	10.5	371	3	AAG45420	Aag45420 Arabidops
35	231.5	10.5	384	2	AAR89411	Aar89411 Stress ac
36	231.5	10.5	426	2	AAR89410	Aar89410 Stress ac
37	231.5	10.5	426	5	ABB85009	Abb85009 Pain regu
38	231.5	10.5	549	8	ADI45327	Adi45327 Rice isop
39	231	10.5	298	2	AAW95689	Aaw95689 Homo sapi
40	231	10.5	298	7	ADD21402	Add21402 Chinese h
41	231	10.5	298	7	ADD21406	Add21406 Rat cycli
42	231	10.5	298	7	ADD21394	Add21394 Mouse cyc
43	231	10.5	416	2	AAW06795	Aaw06795 High Osmo
44	231	10.5	435	6	AAE31910	Aae31910 Yeast Hog
45	231	10.5	435	6	AAE31915	Aae31915 Yeast Hog

ALIGNMENTS

RESULT 1
AAG65767
ID AAG65767 standard; protein; 419 AA.
XX
AC AAG65767;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human protein kinase polypeptide 13887.
XX
KW Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KW cytostatic; antiinflammatory; immunosuppressive; cardiant; hepatotrophic;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WO200173050-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-US009483.
XX
PR 24-MAR-2000; 2000US-0191846P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers R;
XX
DR WPI; 2001-611632/70.
DR N-PSDB; AAI66828, AAI66829.
XX
PT New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT useful in diagnosis of cancer or cellular proliferation or
PT differentiation disorders and to screen for polypeptide modulators useful
PT to treat such conditions.
XX
PS Claim 4; Fig 10A-C; 169pp; English.
XX
CC The invention provides novel human protein kinase polypeptides, 3714,
CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC protein kinase polypeptides can be expressed by standard recombinant
CC methodology.3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC are useful for diagnostic and screening methods to identify subjects (at
CC risk of) having cancer or cellular proliferation and/or differentiation
CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC modulators are useful for the treatment of cancer, particularly colon
CC cancer or cellular proliferation and/or differentiation disorders. Other
CC disorders associated with 3714, 16742, 23546 or 13887 expression or

```
CC   activity that can be treated include bone related disorders, inflammatory
CC   disorders, autoimmune diseases, cardiovascular disorders and liver
CC   diseases. The present sequence represents a human protein kinase
CC   polypeptide 13887
XX
SQ   Sequence 419 AA;

  Query Match             99.5%;  Score 2195;  DB 4;  Length 419;
  Best Local Similarity   99.5%;  Pred. No. 5.3e-229;
  Matches  417;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEEEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 2
ABP64797
ID   ABP64797 standard; protein; 419 AA.
XX
AC   ABP64797;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Human protein SEQ ID 457.
XX
KW   Human; expressed sequence tag; EST; haematopoietic disorder;
KW   central nervous system disease; viral infection;
KW   peripheral nervous system disease; non-healing wound; infectious disease;
KW   immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW   fungal infection; autoimmune disorder; coagulation disorder; nootropic;
KW   antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW   cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW   immunostimulant; cerebroprotective.
XX
OS   Homo sapiens.
XX
PN   WO200259260-A2.
XX
PD   01-AUG-2002.
XX
PF   16-NOV-2001; 2001WO-US042950.
XX
PR   17-NOV-2000; 2000US-00714936.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Goodrich RW,  Liu C,  Zhou P,  Asundi V,  Zhang J,  Zhao QA;
PI   Ren F,  Xue AJ,  Yang Y,  Wehrman T,  Drmanac RT;
XX
DR   WPI; 2002-590824/63.
DR   N-PSDB; ABQ99383.
XX
PT   New isolated polynucleotide, useful in research, diagnostic or
PT   therapeutic methods, e.g. preventing or treating disorders involving
PT   aberrant protein expression or biological activity.
XX
PS   Claim 20; SEQ ID NO 457; 394pp; English.
XX
CC   The present invention relates to novel human coding sequences (ABQ99268-
CC   ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC   therapeutic, diagnostic and research methods. The polynucleotides may be
CC   used in the field of molecular biology as hybridisation probes, primers
CC   for PCR, for chromosome and gene mapping, for the recombinant production
CC   of protein, or in generation of anti-sense DNA or RNA. The
CC   polynucleotides are useful in diagnostics as expressed sequence tags
CC   (ESTs) for identifying expressed genes or for physical mapping of the
CC   human genome. The proteins may be used as molecular weight markers, or as
CC   nutritional sources or supplements. The proteins may be used to maintain
CC   and expand cell population in a totipotential or pluripotential state
CC   useful for re-engineering damaged or diseased tissues, transplantation,
CC   manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC   polynucleotides and proteins are useful for preventing, treating or
CC   ameliorating disorders involving aberrant protein expression or
CC   biological activity, e.g. haematopoietic disorders, central/peripheral
CC   nervous system diseases, mechanical and traumatic disorders, non-healing
CC   wounds, immune deficiencies and disorders, infectious diseases caused by
CC   viral, bacterial or fungal infection, autoimmune disorders, allergic
CC   reactions and conditions, coagulation disorders, or cancer. The
CC   polynucleotide sequences of the invention were assembled from ESTs
CC   isolated mainly by sequencing by hybridisation, and in some cases,
CC   sequences obtained from one or more public databases. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 419 AA;

  Query Match             99.5%;  Score 2195;  DB 5;  Length 419;
  Best Local Similarity   99.5%;  Pred. No. 5.3e-229;
  Matches  417;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEEEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 3
```

```
AAY70306
ID   AAY70306 standard; protein; 418 AA.
XX
AC   AAY70306;
XX
DT   06-JUN-2000  (first entry)
XX
DE   Human transdominant mutant serine/threonine kinase KIS (hKIS) protein.
XX
KW   KIS; human; hKIS; serine/threonine kinase; transdominant mutant; p27;
KW   G1 phase; cell proliferation; cyclin-dependent kinase inhibitor; CKI;
KW   modulator; treatment; cell proliferative disease; vascular disorder;
KW   gene therapy; atherosclerosis; restenosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 54
FT                   /note= "Wild type Lys substituted with Arg"
FT   Misc-difference 186
FT                   /note= "Encoded by AATCAG"
XX
PN   WO200011165-A1.
XX
PD   02-MAR-2000.
XX
PF   20-AUG-1999;   99WO-US018903.
XX
PR   21-AUG-1998;   98US-0097710P.
XX
PA   (NABE/) NABEL G J.
PA   (NABE/) NABEL E G.
XX
PI   Nabel GJ,  Nabel EG;
XX
DR   WPI; 2000-237648/20.
DR   N-PSDB; AAZ51356.
XX
PT   Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT   for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT   proliferation.
XX
PS   Example 1; Page 59-61; 70pp; English.
XX
CC   The present protein sequence is the transdominant mutant human KIS
CC   (hKIS), constructed by site directed mutagenesis. A single nucleotide
CC   substitution (A to G) results in a lysine to arginine change in the
CC   protein sequence. hKIS is a serine/threonine kinase, that acts as an
CC   inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
CC   controls cell proliferation and is localised predominantly in the
CC   nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC   G1 phase. The hKIS sequences are used to modulate cell proliferation and
CC   treat cell proliferative and vascular diseases. The polynucleotide
CC   sequence may be used in gene therapy to treat vascular disorders such as
CC   restenosis or atherosclerosis
XX
SQ   Sequence 418 AA;

  Query Match             99.3%;  Score 2190.5;  DB 3;  Length 418;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-228;
  Matches  418;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGN-DVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 239

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 299

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 359

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 418


RESULT 4
AAU03534
ID   AAU03534 standard; protein; 419 AA.
XX
AC   AAU03534;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Human protein kinase #34.
XX
KW   Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW   metabolic disorder; immune related disease; neurological disorder;
KW   neurodegenerative disorder; inflammatory disorder; infectious disease;
KW   reproductive disorder.
XX
OS   Homo sapiens.
XX
PN   WO200138503-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US032085.
XX
PR   24-NOV-1999;   99US-0167482P.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman GD,  Whyte D,  Manning G,  Sudarsanam S,  Martinez R;
PI   Flanagan P,  Clary D;
XX
DR   WPI; 2001-343950/36.
DR   N-PSDB; AAS06734.
XX
PT   Nucleic acids encoding human kinase polypeptides, useful for preventing
PT   diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT   neuronal-associated diseases, and microbial infections.
XX
PS   Claim 7; Fig 2; 433pp; English.
XX
CC   AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel
CC   protein kinases have been identified as members of the tyrosine or
CC   serine/threonine kinase (PTK and STK) families. The polynucleotides
CC   encoding protein kinases and the polypeptides may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate kinase expression. For example, they may be used to treat
CC   cancers (especially cancers of haematopoietic origin), cardiovascular
CC   disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC   immune related diseases (e.g. rheumatoid arthritis), neurological
CC   disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC   Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC   disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC   Additionally, polynucleotides encoding protein kinases may be used for
CC   gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC   polypeptides may be used as antigens in the production of antibodies
CC   against the protein kinases and in assays to identify modulators of
```

```
CC   protein kinase expression and activity
XX
SQ   Sequence 419 AA;

  Query Match             99.2%;  Score 2188;  DB 4;  Length 419;
  Best Local Similarity   99.3%;  Pred. No. 3.1e-228;
  Matches  416;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEEEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKGNPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 5
AAE11769
ID   AAE11769 standard; protein; 419 AA.
XX
AC   AAE11769;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human kinase (PKIN)-3 protein.
XX
KW   Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
KW   cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
KW   AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
KW   atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
KW   myasthenia gravis; cirrhosis; cataract; growth and development disorder;
KW   seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW   lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW   asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
KW   antimicrobial; cardiovascular; antiinflammatory; vaccine.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          23. .304
FT                   /note= "Eukaryotic protein kinase domain"
FT   Domain          48. .289
FT                   /note= "Protein kinase domain"
FT   Domain          74. .293
FT                   /note= "Protein kinase domain"
FT   Domain          108. .241
FT                   /note= "Protein kinase domain"
FT   Domain          148. .166
FT                   /note= "Tyrosine kinase catalytic domain"
FT   Domain          223. .245
FT                   /note= "Tyrosine kinase catalytic domain"
FT   Domain          223. .241
FT                   /label= Transmembrane_domain
FT   Domain          273. .295
FT                   /note= "Tyrosine kinase catalytic domain"
FT   Region          305. .344
FT                   /note= "Serine/threonine protein kinase"
FT   Region          319. .404
FT                   /note= "Ribonucleoprotein repeat"
FT   Region          320. .412
FT                   /note= "Splicing factor like protein"
XX
PN   WO200181555-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-US012992.
XX
PR   20-APR-2000; 2000US-0199021P.
PR   28-APR-2000; 2000US-0200226P.
PR   05-MAY-2000; 2000US-0202339P.
PR   11-MAY-2000; 2000US-0203505P.
PR   18-MAY-2000; 2000US-0205564P.
PR   26-MAY-2000; 2000US-0207739P.
PR   01-JUN-2000; 2000US-0208795P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Gandhi AR,  Tribouley CM,  Kearney L,  Griffin JA,  Nguyen DB;
PI   Bandman O,  Lu DAM,  Lal P,  Burford N,  Khan FA,  Walia NK,  Yao MG;
PI   Patterson C,  Burrill JD,  Marcus GA,  Zingler KA,  Recipon SA,  Lu Y;
PI   Policky JL,  Thornton M,  Tang YT,  Hafalia A,  Elliott VS,  Baughn MR;
PI   Walsh RT,  Ramkumar J,  Borowsky ML,  Au-Young J,  Hillman JL;
PI   Gururajan R;
XX
DR   WPI; 2001-611740/70.
DR   N-PSDB; AAD18818.
XX
PT   Human kinases and nucleic acids, useful for preventing diagnosing and
PT   treating cancers, inflammation and immune disorders.
XX
PS   Claim 1; Page 125-126; 166pp; English.
XX
CC   The present invention relates to human kinases (PKIN) and the nucleic
CC   acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
CC   used in the prevention, diagnosis and treatment of diseases cancers,
CC   adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
CC   acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
CC   gout, microbial infections, cardiovascular disease and/or inflammation,
CC   myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
CC   infarction, cataract, growth and development disorder, seizure disorder,
CC   pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
CC   disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
CC   PKIN may be used to treat disorders associated with decreased PKIN
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of PKIN by expressing inactive proteins or to
CC   supplement the patients own production of PKIN. PKIN nucleic acids may be
CC   used to produce the PKIN polypeptide, by inserting the nucleic acids into
CC   a host cell and culturing the cell to express the protein. PKIN nucleic
CC   acid and its complementary sequences may also be used as DNA probes in
CC   diagnostic assays to detect and quantitate the presence of similar
CC   nucleic acid sequences in samples and therefore which patients may be in
CC   need of restorative therapy. The present sequence is human PKIN-3 protein
XX
SQ   Sequence 419 AA;

  Query Match             99.2%;  Score 2188;  DB 4;  Length 419;
  Best Local Similarity   99.3%;  Pred. No. 3.1e-228;
  Matches  416;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy         121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy         181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy         241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy         301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
                |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEEEYEDVVEDVKEECQKYGPVVSL 360

Qy         361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
                |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 LVPKGNPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
```

RESULT 6
AAY70305
ID AAY70305 standard; protein; 418 AA.
XX
AC AAY70305;
XX
DT 06-JUN-2000 (first entry)
XX
DE Human wild type serine/threonine kinase KIS (hKIS) protein.
XX
KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
KW cyclin-dependent kinase inhibitor; CKI; p27; modulator; treatment;
KW cell proliferative disease; vascular disorder; gene therapy; restenosis;
KW atherosclerosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 186
FT /note= "Encoded by AATCAG"
XX
PN WO200011165-A1.
XX
PD 02-MAR-2000.
XX
PF 20-AUG-1999; 99WO-US018903.
XX
PR 21-AUG-1998; 98US-0097710P.
XX
PA (NABE/) NABEL G J.
PA (NABE/) NABEL E G.
XX
PI Nabel GJ, Nabel EG;
XX
DR WPI; 2000-237648/20.
DR N-PSDB; AAZ51355.
XX
PT Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT proliferation.
XX
PS Example 1; Page 57-58; 70pp; English.
XX
CC The present protein sequence is the wild type human KIS (hKIS), isolated
CC from a human B-cell library, using a yeast two hybrid screening system.
CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS controls cell
CC proliferation and is localised predominantly in the nucleus. It binds to
CC CKI p27 and inhibits its ability to arrest cells in G1 phase. It has an
CC amino terminal serine/threonine kinase consensus region and a carboxy
CC terminus with 42% sequence similarity to hU2AF65, a subunit of the
CC splicing factor U2AF. The hKIS sequences are used to modulate cell
CC proliferation and treat cell proliferative and vascular diseases. The DNA
CC sequence may be used in gene therapy to treat vascular disorders such as
CC restenosis or atherosclerosis
XX
SQ Sequence 418 AA;

Query Match 99.2%; Score 2187.5; DB 3; Length 418;
Best Local Similarity 99.5%; Pred. No. 3.5e-228;
Matches 417; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```
Qy           1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db           1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALKQFLPPG 60

Qy          61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy         121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy         181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
                |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 SFKEGN-DVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 239

Qy         241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         240 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 299

Qy         301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         300 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 359

Qy         361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         360 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 418
```

RESULT 7
ADD48093
ID ADD48093 standard; protein; 419 AA.
XX
AC ADD48093;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein AAC53031, SEQ ID NO 13791.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.

```
DR   GENBANK; AAC53031.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a rat protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 419 AA;

  Query Match             99.0%;  Score 2183;  DB 7;  Length 419;
  Best Local Similarity   98.8%;  Pred. No. 1.1e-227;
  Matches  414;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGTPGSPPGALKQFLPPG 60

Qy         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE 120

Qy        121 LLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGL 180

Qy        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMK 240

Qy        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYHLRDLIKSMLHDDPSRRIPAEMALC 300

Qy        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEECQKYGPVVSL 360
              |||||||||||||||||||||||||||||||| ||:||||||||||||||||||||||||
Db        301 SPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLENEDEYEDVVEDVKEECQKYGPVVSL 360

Qy        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVATFYPLSAYKRGYLYQTLL 419


RESULT 8
ABP64662
ID   ABP64662 standard; protein; 137 AA.
XX
AC   ABP64662;
XX
DT   04-NOV-2002  (first entry)
XX
DE   Human ORF1032.
XX
KW   Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW   Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW   human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW   cancer; cardiovascular disease; allergy; autoimmune disease;
KW   wound healing; blood coagulation disorder; inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   US2002082206-A1.
XX
PD   27-JUN-2002.
XX
PF   30-MAY-2001; 2001US-00867550.
XX
PR   30-MAY-2000; 2000US-0208427P.
XX
PA   (LEAC/) LEACH M D.
PA   (MEHR/) MEHRABAN F.
PA   (CONL/) CONLEY P B.
PA   (TOPP/) TOPPER J N.
PA   (LAWD/) LAW D.
XX
PI   Leach MD,  Mehraban F,  Conley PB,  Topper JN,  Law D;
XX
DR   WPI; 2002-626554/67.
DR   N-PSDB; ABQ99225.
XX
PT   New polypeptide designated ORFX are present in human atherogenic cells
PT   and are useful to prevent and treat ORFX-associated disorders including
PT   cancer, allergy, wound healing or autoimmune, cardiovascular or
PT   inflammatory disease.
XX
PS   Claim 10; SEQ ID NO 2064; 78pp; English.
XX
CC   The present invention relates to novel human ORFX polypeptides and their
CC   coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC   were discovered in human atherogenic cells, in particular in platelets
CC   and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC   many other tissues as well. Atherogenic cells are cells which have the
CC   potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC   nucleic acids are useful for treating or preventing a pathological
CC   condition associated with an ORFX-associated disorder, e.g. cancer,
CC   cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC   coagulation disorders or inflammatory disorders. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/sequence.html?DocID=20020082206
XX
SQ   Sequence 137 AA;

  Query Match             21.9%;  Score 483.5;  DB 5;  Length 137;
  Best Local Similarity   80.0%;  Pred. No. 1e-43;
  Matches   96;  Conservative    1;  Mismatches   12;  Indels   11;  Gaps    1;

Qy         63 GAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELL 122
              |  |:|   | |   |           ||||||||||||||||||||||||||||||
Db         29 GRGAAAFSGGLRAAVA----------VTLYGVFTIHFSPNVPSRCLLLELLDVSVSELL 77

Qy        123 LYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSF 182
              |||||||||||||||||||||||||||||||||||||| ||||||||||| ||||||||
Db         78 LYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLXPRNILWSAENECXTLIDFGLSF 137


RESULT 9
AAG38502
ID   AAG38502 standard; protein; 638 AA.
```

XX
AC AAG38502;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47508.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.

```
PR   25-AUG-1999;   99US-0150566P.
PR   26-AUG-1999;   99US-0150884P.
PR   27-AUG-1999;   99US-0151065P.
PR   27-AUG-1999;   99US-0151066P.
PR   27-AUG-1999;   99US-0151080P.
PR   30-AUG-1999;   99US-0151303P.
PR   31-AUG-1999;   99US-0151438P.
PR   01-SEP-1999;   99US-0151930P.
PR   07-SEP-1999;   99US-0152363P.
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match             10.9%;  Score 240.5;  DB 3;  Length 638;
  Best Local Similarity   25.7%;  Pred. No. 3e-16;
  Matches  105;  Conservative   57;  Mismatches  146;  Indels  101;  Gaps   17;

Qy         11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPGTTGAAASAAE 70
              :|  | :  |:|:|   ::||| |: |::|    :      ||::    |   | |
Db        249 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 301

Qy         71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
              ||| :|   |::|:|  ||: |            |  : || |:|: ::|
Db        302 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 354

Qy        121 ---LLLYS----SHQGCSMW------------MIQHCARDVLEALAFLHHEGYVHADLK 160
                :| |     :|     |             ::    : :|:|:  :| |  ||:|||
Db        355 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 414

Qy        161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
              | | |        ||||||:: |  | |   || |       | :||| : |   : |
Db        415 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 469

Qy        215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272
              :  :|    |:|||| || :|  | :          | |    | | |      : |
Db        470 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 525

Qy        273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH---------------------- 311
                 | ||:|  |  | ::|       |  || : |  ||
Db        526 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 581

Qy        312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351
                      :| |:  | |: |      :| | : :: : ::  | | |
Db        582 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 624


RESULT 10
AAG38501
ID   AAG38501 standard; protein; 720 AA.
XX
AC   AAG38501;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 47507.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
```

PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 10.9%; Score 240.5; DB 3; Length 720;
Best Local Similarity 25.7%; Pred. No. 3.7e-16;

```
  Matches  105;  Conservative   57;  Mismatches  146;  Indels  101;  Gaps   17;

Qy          11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPGTTGAAASAAE 70
               :|  | :  |:|:|   ::||| |: |::|    :      ||::    |   | |
Db         331 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 383

Qy          71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
               ||| :|   |::|:|  ||: |           |  : || |:|: ::|
Db         384 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 436

Qy         121 ---LLLYS----SHQGCSMW------------MIQHCARDVLEALAFLHHEGYVHADLK 160
                  :| |     :|     |              ::   : :|:|:  :| |  ||:|||
Db         437 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 496

Qy         161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
               | | |        ||||||:: |  | |   || |       | :||| : |   : |
Db         497 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 551

Qy         215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272
               :  :|    |:|||| || :|  | :         | |    | |  |      : |
Db         552 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 607

Qy         273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH--------------------- 311
                   | ||:|  |  | ::|      |  || : |  ||
Db         608 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 663

Qy         312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351
                       :| |:  | |: |      :| | : :: : ::  | | |
Db         664 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 706


RESULT 11
AAG38500
ID   AAG38500 standard; protein; 777 AA.
XX
AC   AAG38500;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 47506.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   20-JUL-1999;   99US-0144884P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
```

```
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
PR   23-JUL-1999;   99US-0145224P.
PR   26-JUL-1999;   99US-0145276P.
PR   27-JUL-1999;   99US-0145913P.
PR   27-JUL-1999;   99US-0145918P.
PR   27-JUL-1999;   99US-0145919P.
PR   28-JUL-1999;   99US-0145951P.
PR   02-AUG-1999;   99US-0146386P.
PR   02-AUG-1999;   99US-0146388P.
PR   02-AUG-1999;   99US-0146389P.
PR   03-AUG-1999;   99US-0147038P.
PR   04-AUG-1999;   99US-0147204P.
PR   04-AUG-1999;   99US-0147302P.
PR   05-AUG-1999;   99US-0147192P.
PR   05-AUG-1999;   99US-0147260P.
PR   06-AUG-1999;   99US-0147303P.
PR   06-AUG-1999;   99US-0147416P.
PR   09-AUG-1999;   99US-0147493P.
PR   09-AUG-1999;   99US-0147935P.
PR   10-AUG-1999;   99US-0148171P.
PR   11-AUG-1999;   99US-0148319P.
PR   12-AUG-1999;   99US-0148341P.
PR   13-AUG-1999;   99US-0148565P.
PR   13-AUG-1999;   99US-0148684P.
PR   16-AUG-1999;   99US-0149368P.
PR   17-AUG-1999;   99US-0149175P.
PR   18-AUG-1999;   99US-0149426P.
PR   20-AUG-1999;   99US-0149722P.
PR   20-AUG-1999;   99US-0149723P.
PR   20-AUG-1999;   99US-0149929P.
PR   23-AUG-1999;   99US-0149902P.
PR   23-AUG-1999;   99US-0149930P.
PR   25-AUG-1999;   99US-0150566P.
PR   26-AUG-1999;   99US-0150884P.
PR   27-AUG-1999;   99US-0151065P.
PR   27-AUG-1999;   99US-0151066P.
PR   27-AUG-1999;   99US-0151080P.
PR   30-AUG-1999;   99US-0151303P.
PR   31-AUG-1999;   99US-0151438P.
PR   01-SEP-1999;   99US-0151930P.
PR   07-SEP-1999;   99US-0152363P.
PR   10-SEP-1999;   99US-0153070P.
PR   13-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match           10.9%;  Score 240.5;  DB 3;  Length 777;
  Best Local Similarity   25.7%;  Pred. No. 4.2e-16;
  Matches  105;  Conservative   57;  Mismatches  146;  Indels  101;  Gaps   17;

Qy         11 EPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQFLPPGTTGAAASAAE 70
              :|  | :  |:|:|   ::||| |: |::|    :       ||::    |   | |
Db        388 DPDLFFKVNGKLYQRLGKIGSGGSSEVHKVI---SSDCTIYALKKIKLKGRDYATA---- 440

Qy         71 YGFRKERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSE---------- 120
              ||| :|   |::|:|  ||: |           |  : || |:|: ::|
Db        441 YGFCQEIGYLKKLKGKTNIIQL-------IDYEVTDKTLLQEVLNGTMSNKDGRVKEDGF 493

Qy        121 ---LLLYS----SHQGCSMW------------MIQHCARDVLEALAFLHHEGYVHADLK 160
                :| |     :|     |               ::   : :|:|:  :| |  ||:|||
Db        494 IYMVLEYGEIDLAHMLSQKWREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLK 553

Qy        161 PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTD------GYRAPEAELQNCLAQAGLQ 214
              | | |        ||||||:: |  | |   || |       | :||| : |   : |
Db        554 PANFL--LVRGFLKLIDFGIA-KAINSDTTNIQRDSQVGTLSYMSPEAFMCNESDENG-- 608

Qy        215 SDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIID--HIFASKAVVNAAI 272
              :  :|    |:|||| || :|  | :          | |    | |  |      : |
Db        609 NTIKCGRPSDIWSLGCILYQMVYG-RTPFADYKTFW-AKFKVITDPNHEITYNQLSNP-- 664

Qy        273 PAYHLRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPH---------------------- 311
                | ||:|  |  | ::|       |  || : |  ||
Db        665 ---WLIDLMKKCLAWDRNQRWRIPELLQHPFLAPPI-PHEPQVKTIKLFSLIAESCGSDD 720

Qy        312 ---------IEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDVKEEC 351
                      :| |:  | |: |      :| | : :: : ::  | | |
Db        721 DKANSMISQLEQLLSNPAPLPR------NDVLDSRDQNQQLLSRVSELC 763


RESULT 12
AAW85028
ID   AAW85028 standard; protein; 544 AA.
XX
AC   AAW85028;
XX
DT   17-OCT-2003  (revised)
DT   08-FEB-1999  (first entry)
XX
DE   CDK2-green fluorescent protein usion product.
XX
KW   Human; CDK2 gene; fusion protein; green fluorescent protein; GFP;
KW   intracellular signalling; chimera.
XX
OS   Aequorea victoria.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   WO9845704-A2.
XX
PD   15-OCT-1998.
XX
PF   07-APR-1998;   98WO-DK000145.
XX
PR   07-APR-1997;   97DK-00000392.
```

```
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Thastrup O,  Petersen Bjorn S,  Tullin S,  Kasper A,  Scudder K;
XX
DR   WPI; 1998-594491/50.
DR   N-PSDB; AAV71073.
XX
PT   Determining effect on signalling pathways in live cells from
PT   redistribution of luminophores - specifically fusions of green
PT   fluorescent protein with a signalling component, and new apparatus,
PT   particularly for identifying toxins and potential therapeutic agents.
XX
PS   Example 12; Page 208-209; 326pp; English.
XX
CC   The present sequence represents a human CDK2-green fluorescent protein
CC   (GFP) fusion protein. The fusion protein is used in an assay to exemplify
CC   the invention. The specification describes how quantitative information
CC   about the influence of a molecule on a cellular response is obtained by
CC   recording the variation, caused by the molecule, on mechanically intact
CC   living cells, in the spatially distributed light emitted from a
CC   luminophore present in the cells. The variation in light emission is
CC   processed to provide information that correlates spatial distribution to
CC   the degree of the molecule. The method is used to identify agents that
CC   (in)directly affect intracellular signalling, especially to screen for
CC   potential therapeutic agents or toxins, and to identify new drug targets.
CC   (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 544 AA;

  Query Match             10.7%;  Score 236.5;  DB 2;  Length 544;
  Best Local Similarity   26.7%;  Pred. No. 6.5e-16;
  Matches  111;  Conservative   53;  Mismatches  154;  Indels   97;  Gaps   18;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQF-LPPGTTGAAASAAEYGFRKERAALE 81
              :|   ::| |:   ||: |          ||::  |   | |  ::|      :| : |:
Db          4 FQKVEKIGEGTYGVVYKAR--NKLTGEVVALKKIRLDTETEGVPSTAI-----REISLLK 56

Qy         82 QLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARD 141
              :|  | ||| |  |  ||    :    |: | |   : : :  |:  |  : :|:
Db         57 EL-NHPNIVKLLDV--IHTENKL---YLVFEFLHQDLKKFMDASALTGIPLPLIKSYLFQ 110

Qy        142 VLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY---IQTDGYR 198
              :|: ||| |    :| ||||:|:| : |    || ||||:   |     |   : |  ||
Db        111 LLQGLAFCHSHRVLHRDLKPQNLLINTEG-AIKLADFGLARAFGVPVRTYTHEVVTLWYR 169

Qy        199 APEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAII 258
              |||           |      ::|||:|||| |  ||          |   :  :|   |
Db        170 APEI----------LLGSKYYSTAVDIWSLGCIFAEM--------VTRRALFPGDSE--I 209

Qy        259 DHIFASKAVVN----------AAIPAYHL--------------------RDLIKSMLHD 287
              | :|     :           ::| |                      | |:  |||
Db        210 DQLFRIFRTLGTPDEVVWPGVTSMPDYKPSFPKWARQDFSKVVPPLDEDGRSLLSQMLHY 269

Qy        288 DPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYLGNEEEYEDVVEDV 347
              ||::|| |: ||  |||         :| |  | | ||| :             |   |
Db        270 DPNKRISAKAALAHPFF--------QD-VTKPVPHLRLWD-------------PPVATMV 307

Qy        348 KEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGRMFDGKFVVAT 402
               :  : :  || :||  :    |  |   :  |:  |    ||      ||:  |
Db        308 SKGEELFTGVVPILVELDGDVNGHKF-SVSGEGEGDATYGKLT-----LKFICTT 356


RESULT 13
ADD21398
ID   ADD21398 standard; protein; 298 AA.
XX
AC   ADD21398;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Golden hamster cyclin dependent kinase (cdk) 2 protein.
XX
KW   continual growth; cultured cell; cyclin dependent kinase; cdk4; cdk2;
KW   cdk6; activating mutation; cell growth; cell division; cell cycle;
KW   cancer-causing agent; continual growth-induced cell; enzyme;
KW   golden hamster.
XX
OS   Mesocricetus auratus.
XX
PN   WO2003044169-A2.
XX
PD   30-MAY-2003.
XX
PF   15-NOV-2002; 2002WO-US036729.
XX
PR   15-NOV-2001; 2001US-0334760P.
XX
PA   (UTEM ) UNIV TEMPLE.
XX
PI   Reddy PE,  Rane SG,  Mettus RV;
XX
DR   WPI; 2003-449813/42.
DR   N-PSDB; ADD21399.
XX
PT   A composition for reversibly inducing continual growth in normal cells
PT   comprises a cyclin dependent kinase protein (e.g. cdk4, cdk2 or cdk6) or
PT   its active fragment, derivative, homolog or analog, having an activating
PT   mutation.
XX
PS   Disclosure; Page 104; 77pp; English.
XX
CC   This invention relates to a novel composition for inducing a reversible
CC   state of a continual growth in cultured cells and comprises at least one
CC   compound comprising a cyclin dependent kinase (cdk)4, cdk2 or cdk6
CC   protein having an activating mutation. Growth and division of living
CC   cells involve a regular series of events and processes that comprise the
CC   cell cycle. Cyclin dependent kinases cdk2, cdk4 and cdk6 are involved in
CC   the control of G1, the point at which cells irrevocably commit to DNA
CC   synthesis and thus enter the cell cycle. The invention is useful in
CC   reversibly inducing continual growth in normal cells and may allow the
CC   screening of cancer-causing agents with the continual growth-induced
CC   cells. The present sequence is that of the golden hamster cdk2 protein
CC   related to the invention. Note: Due to an error in the specification or
CC   sequence listing, the Seq ID numbers given in the disclosure do not
CC   correspond to those given in the sequence listing. It is therefore
CC   unclear which Seq ID number corresponds to which sequence and exactly
CC   which sequence is being claimed.
XX
SQ   Sequence 298 AA;

  Query Match             10.7%;  Score 235;  DB 7;  Length 298;
  Best Local Similarity   29.0%;  Pred. No. 3.7e-16;
  Matches   98;  Conservative   45;  Mismatches  117;  Indels   78;  Gaps   16;

Qy         23 WQVQSRLGSGSSASVYRVRCCGNPGSPPGALRQF-LPPGTTGAAASAAEYGFRKERAALE 81
              :|   ::| |:   ||: :          ||::  |   | |  ::|      :| : |:
Db          4 FQKVEKIGEGTYGVVYKAK--NKLTGEVVALKKIRLDTETEGVPSTAI-----REISLLK 56

Qy         82 QLQGHRNIVTLYGVFTIHFSPNVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARD 141
              :|  | ||| |  |  ||    :    |: |||   : : :  |:  |  : :|:
Db         57 EL-NHPNIVKLLDV--IHTENKL---YLVFELLHQDLKKFMDASAVTGIPLPLIKSYLFQ 110

Qy        142 VLEALAFLHHEGYVHADLKPRNILWSAENECFKLIDFGLSFKEGNQDVKY---IQTDGYR 198
              :|: ||| |    :| ||||:|:| :||    || ||||:   |     |   : |  ||
Db        111 LLQGLAFCHSHRVLHRDLKPQNLLINAEGS-IKLADFGLARAFGVPVRTYTHEVVTLWYR 169

Qy        199 APEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAII 258
              |||  |  |            ::|||:|||| |  ||          |   :  :|   |
Db        170 APEI-LLGC---------KYYSTAVDIWSLGCIFAEM--------VTRRALFPGDSE--I 209

Qy        259 DHIFASKAVVN----------AAIPAYHL--------------------RDLIKSMLHD 287
              | :|     :           ::| |                      | |:  |||
Db        210 DQLFRIFRTLGTPDEVVWPGVTSMPDYKPSFPKWARQDFSKVVPPLDEDGRSLLSQMLHY 269

Qy        288 DPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRL 325
```

```
                ||::|| |: ||  |||        :| |  | | |||
Db        270 DPNKRISAKAALAHPFF--------QD-VTKPVPHLRL 298


RESULT 14
AAB19094
ID   AAB19094 standard; protein; 433 AA.
XX
AC   AAB19094;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Protein encoded by reporter gene YKL161C from protein kinase C pathway.
XX
KW   Reporter gene; ergosterol pathway; protein kinase C pathway;
KW   invasive growth pathway; biological pathway.
XX
OS   Saccharomyces cerevisiae.
XX
PN   WO200058520-A1.
XX
PD   05-OCT-2000.
XX
PF   29-MAR-2000; 2000WO-US008555.
XX
PR   31-MAR-1999;   99US-00282243.
XX
PA   (ROSE-) ROSETTA INPHARMATICS INC.
XX
PI   Roberts CJ;
XX
DR   WPI; 2000-664929/64.
DR   N-PSDB; AAA96704.
XX
PT   Identifying reporter and target genes for particular biological pathway
PT   of interest, useful for drug designing, involves identifying a gene which
PT   clusters to geneset associated with biological pathway.
XX
PS   Disclosure; Fig 20; 239pp; English.
XX
CC   The present sequence is encoded by a reporter gene from the protein
CC   kinase C pathway. The specification also describes reporter genes from
CC   the ergosterol pathway, and the invasive growth pathway. The genes are
CC   identified using methods of the invention. The specification describes
CC   methods for identifying a reporter gene or a target gene for a particular
CC   biological pathway in a cell. The methods comprise identifying a gene
CC   which clusters to a geneset associated with the pathway. The gene
CC   identified is necessary for normal function of the pathway. The methods
CC   are useful for analysing the activity of biological pathways and in drug
CC   designing, drug therapies, or other biological agents e.g. insecticides,
CC   herbicides, fungicides, antibodies or antivirals. The reporter gene is
CC   useful as a specific target for drugs which are designed to enhance,
CC   inhibit or modulate a particular biological pathway
XX
SQ   Sequence 433 AA;

  Query Match             10.6%;  Score 234.5;  DB 3;  Length 433;
  Best Local Similarity   29.5%;  Pred. No. 7.4e-16;
  Matches   89;  Conservative   45;  Mismatches   97;  Indels   71;  Gaps   16;

Qy         75 KERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLL--ELLDVSVSELLLYSSHQGCSM 132
              :|   |  |:|| ||| |:    | | ||     : |  ||::  :|:::   | |
Db         71 RELKLLRHLRGHPNIVWLFDT-DIVFYPNGALNGVYLYEELMECDLSQII--RSEQRLED 127

Qy        133 WMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEG 185
                  |     :| || ::|    :| ||||:|:|    ::|  |: :||||      |
Db        128 AHFQSFIYQILCALKYIHSANVLHCDLKPKNLL--VNSDCQLKICNFGLSCSYSENHKVN 185

Qy        186 NQDVK-YIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLE------MFSG 238
              :  :| || :  |:|||           | :  ||| |||:|| | || |      || |
Db        186 DGFIKGYITSIWYKAPEI----------LLNYQECTKAVDIWSTGCILAELLGRKPMFEG 235

Qy        239 M----KLKHTVR---------SQEWKANSSAIIDHIF--------ASKAVVNAAIPAYHL 277
                  | | ::          ||    |  : ::||        : ::::  | |
Db        236 KDYVDHLNHILQILGTPPEETLQE--IASQKVYNYIFQFGNIPGRSFESILPGANP--EA 291

Qy        278 RDLIKSMLHDDPSRRIPAEMALCSPFFSI------PFA---------PHIEDLVMLPTPV 322
               :|:| ||  || :||  | ||  |: |:       |:          ||| :  |   |
Db        292 LELLKKMLEFDPKKRITVEDALEHPYLSMWHDIDEEFSCQKTFRFEFEHIESMAELGNEV 351

Qy        323 LR 324
              ::
Db        352 IK 353


RESULT 15
ADP87472
ID   ADP87472 standard; protein; 433 AA.
XX
AC   ADP87472;
XX
DT   23-SEP-2004  (first entry)
XX
DE   S cerevisiae glucan synthase pathway protein YKL161C SeqID14.
XX
KW   glucan synthase pathway; RNA expression; protein expression; YOL1 13W;
KW   SKM1; YNR066C; YLR121C; YPS3; YHR209W; YKL161C; YFR030W; FLO1; YFR030W;
KW   MET1O; antifungal agent.
XX
OS   Saccharomyces cerevisiae.
XX
PN   WO2004057033-A1.
XX
PD   08-JUL-2004.
XX
PF   17-DEC-2003; 2003WO-US040532.
XX
PR   19-DEC-2002; 2002US-00324035.
XX
PA   (ROSE-) ROSETTA INPHARMATICS LLC.
XX
PI   Phillips JW;
XX
DR   WPI; 2004-500308/47.
DR   N-PSDB; ADP87471.
XX
PT   Determining whether a molecule affects the function or activity of a
PT   glucan synthase pathway in a S. cerevisiae cell by determining a change
PT   in the RNA expression or protein expression in the cell of at least one
PT   target polynucleotide.
XX
PS   Disclosure; SEQ ID NO 14; 132pp; English.
XX
CC   This invention relates to a novel method of determining if a molecule
CC   affects the function or activity of a glucan synthase pathway in a
CC   Saccharomyces cerevisiae cell which comprises determining if the RNA
CC   expression or protein expression in the cell of at least one target
CC   polynucleotide sequence is changed relative to the expression of said
CC   target polynucleotide sequence in the absence of the molecule. The target
CC   polynucleotide sequence is selected from the group consisting of YOL1 13W
CC   (SKM1), YNR066C, YLR121C (YPS3), YHR209W, YKL161C, YFR030W (FLO1) and
CC   YFR030W (MET1O). The method is useful for determining whether a molecule
CC   affects the function or activity of a glucan synthase pathway in an S
CC   cerevisiae cell, possibly allowing development of antifungal agents for
CC   use against a variety of pathogens. The present sequence is that of the
CC   protein encoded by an S cerevisiae gene which may be used in the method
CC   of the invention.
XX
SQ   Sequence 433 AA;

  Query Match             10.6%;  Score 234.5;  DB 8;  Length 433;
  Best Local Similarity   29.5%;  Pred. No. 7.4e-16;
  Matches   89;  Conservative   45;  Mismatches   97;  Indels   71;  Gaps   16;

Qy         75 KERAALEQLQGHRNIVTLYGVFTIHFSPNVPSRCLLL--ELLDVSVSELLLYSSHQGCSM 132
              :|   |  |:|| ||| |:    | | ||     : |  ||::  :|:::   | |
```

```
Db         71 RELKLLRHLRGHPNIVWLFDT-DIVFYPNGALNGVYLYEELMECDLSQII--RSEQRLED 127

Qy        133 WMIQHCARDVLEALAFLHHEGYVHADLKPRNILWSAENEC-FKLIDFGLS------FKEG 185
              |     :| || ::|    :| ||||:|:|    ::|  |: :||||       |
Db        128 AHFQSFIYQILCALKYIHSANVLHCDLKPKNLL--VNSDCQLKICNFGLSCSYSENHKVN 185

Qy        186 NQDVK-YIQTDGYRAPEAELQNCLAQAGLQSDTECTSAVDLWSLGIILLE------MFSG 238
              :  :| || :  |:|||           | :  ||| |||:|| | || |       || |
Db        186 DGFIKGYITSIWYKAPEI----------LLNYQECTKAVDIWSTGCILAELLGRKPMFEG 235

Qy        239 M----KLKHTVR---------SQEWKANSSAIIDHIF--------ASKAVVNAAIPAYHL 277
                   | | ::          ||    |  : ::||        : ::::  | |
Db        236 KDYVDHLNHILQILGTPPEETLQE--IASQKVYNYIFQFGNIPGRSFESILPGANP--EA 291

Qy        278 RDLIKSMLHDDPSRRIPAEMALCSPFFSI------PFA---------PHIEDLVMLPTPV 322
              :|:| ||  || :||  | ||  |: |:       |:          ||| :  |   |
Db        292 LELLKKMLEFDPKKRITVEDALEHPYLSMWHDIDEEFSCQKTFRFEFEHIESMAELGNEV 351

Qy        323 LR 324
              ::
Db        352 IK 353
```

Search completed: November 29, 2004, 11:30:38
Job time : 111.5 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 4563 Seconds
(without alignments)
4342.403 Million cell updates/sec

Title: US-10-798-532-4
Perfect score: 2206
Sequence: 1 MAGSGCAWGAEPPRFLEAFG.....VATFYPLSAYKRGYLYQTLL 419

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10798532/runat_29112004_112452_15890/app_query.fasta_1.1166
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532_@CGN_1_1_6317_@runat_29112004_112452_15890 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2195	99.5	1260	6	CQ715025	CQ715025 Sequence
2	2195	99.5	1260	6	AX262521	AX262521 Sequence
3	2195	99.5	2598	6	AX262519	AX262519 Sequence
4	2195	99.5	2909	9	HSA536197	AJ536197 Homo sapi
5	2195	99.5	4065	6	CQ412057	CQ412057 Sequence
6	2195	99.5	4065	6	AX188296	AX188296 Sequence
7	2188	99.2	1260	6	AX166543	AX166543 Sequence
8	2188	99.2	1824	6	AX680136	AX680136 Sequence
9	2183	99.0	1633	10	RNKISRNA	X98374 R.norvegicu
10	2183	99.0	3244	10	RNU70372	U70372 Rattus norv
11	2182	98.9	2041	10	BC058732	BC058732 Mus muscu
c 12	2174	98.5	110000	2	AC129141_1	Continuation (2 of
13	2157	97.8	1254	10	MMPKKIS	Y10725 Mus musculu
14	1990	90.2	1367	9	BC026046	BC026046 Homo sapi
15	1890.5	85.7	2940	9	HSM805327	AL834136 Homo sapi
16	1761.5	79.9	1157	5	CR386837	CR386837 Gallus ga
17	1721	78.0	1135	9	AK058195	AK058195 Homo sapi
18	1634	74.1	2041	9	BC014917	BC014917 Homo sapi
c 19	1050	47.6	259562	2	AC094499	AC094499 Rattus no
20	1045	47.4	601	10	MMKISRNA	X82320 M.musculus
21	905	41.0	105902	9	AL359699	AL359699 Human DNA
22	905	41.0	167995	2	AC040901	AC040901 Homo sapi
c 23	901	40.8	211812	2	AC027205	AC027205 Homo sapi
24	802.5	36.4	211635	2	AC112763	AC112763 Rattus no
25	789	35.8	529	6	AX186343	AX186343 Sequence
26	766.5	34.7	722	6	BD221413	BD221413 Human gen
27	729	33.0	470	6	AX187546	AX187546 Sequence
28	721	32.7	100207	10	AY180177	AY180177 Mus muscu
c 29	694.5	31.5	200095	10	AC121606	AC121606 Mus muscu
30	694.5	31.5	220806	10	AC123650	AC123650 Mus muscu
31	663	30.1	462	6	AX184680	AX184680 Sequence
32	366	16.6	60640	10	AY180176	AY180176 Mus muscu
33	366	16.6	185810	2	AC136052	AC136052 Rattus no
34	361	16.4	231	6	BD220137	BD220137 Human gen
35	327.5	14.8	205752	2	AC119431	AC119431 Mus muscu
36	311.5	14.1	419	6	CQ393016	CQ393016 Sequence
37	311.5	14.1	419	6	CQ399411	CQ399411 Sequence
c 38	267	12.1	185810	2	AC136052	AC136052 Rattus no
39	254	11.5	110000	8	CR382128_04	Continuation (5 of
40	250.5	11.4	1269	10	AB096078	AB096078 Mus muscu
41	250.5	11.4	1395	10	AB096079	AB096079 Mus muscu
42	249	11.3	1065	10	HAMCDK2A	D17350 Mesocricetu
43	248.5	11.3	2809	3	AY064464	AY064464 Caenorhab
44	246.5	11.2	1400	10	RATCDK2A	D28753 Rattus ratt
45	246.5	11.2	2250	10	BC061832	BC061832 Rattus no

ALIGNMENTS

RESULT 1
CQ715025
LOCUS CQ715025 1260 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 959 from Patent WO02068579.
ACCESSION CQ715025
VERSION CQ715025.1 GI:42275882
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 959 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1. .1260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.: 2.07e-188 Length: 1260
Score: 2195.00 Matches: 417

```
Percent Similarity:     99.76%          Conservative:  1
Best Local Similarity:  99.52%          Mismatches:    1
Query Match:            99.50%          Indels:        0
DB:                     6               Gaps:          0

US-10-798-532-4 (1-419) x CQ715025 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
```

```
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257
```

RESULT 2
AX262521

```
LOCUS       AX262521                1260 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 12 from Patent WO0173050.
ACCESSION   AX262521
VERSION     AX262521.1  GI:16511408
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Meyers,R.
  TITLE     3714, 16742, 23546, and 13887 novel protein kinase molecules and
            uses therefor
  JOURNAL   Patent: WO 0173050-A 12 04-OCT-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1260
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.07e-188       Length:         1260
Score:                  2195.00         Matches:        417
Percent Similarity:     99.76%          Conservative:   1
Best Local Similarity:  99.52%          Mismatches:     1
Query Match:            99.50%          Indels:         0
DB:                     6               Gaps:           0

US-10-798-532-4 (1-419) x AX262521 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257
```

```
RESULT 3
AX262519
LOCUS       AX262519                 2598 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 10 from Patent WO0173050.
ACCESSION   AX262519
VERSION     AX262519.1  GI:16511406
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Meyers,R.
  TITLE     3714, 16742, 23546, and 13887 novel protein kinase molecules and
            uses therefor
  JOURNAL   Patent: WO 0173050-A 10 04-OCT-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2598
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             268. .1527
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD10238.1"
                     /db_xref="GI:16511407"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

Alignment Scores:
Pred. No.:              5.7e-188        Length:         2598
Score:                  2195.00         Matches:        417
Percent Similarity:     99.76%          Conservative:   1
Best Local Similarity:  99.52%          Mismatches:     1
Query Match:            99.50%          Indels:         0
DB:                     6               Gaps:           0

US-10-798-532-4 (1-419) x AX262519 (1-2598)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807
```

```
Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||    ||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1524
```

RESULT 4
HSA536197

```
LOCUS       HSA536197               2909 bp    mRNA    linear   PRI 11-JAN-2003
DEFINITION  Homo sapiens mRNA for KIS protein.
ACCESSION   AJ536197
VERSION     AJ536197.1  GI:27657360
KEYWORDS    KIS protein.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Bieche,I., Manceau,V., Curmi,P.A., Laurendon,I., Lachkar,S.,
            Leroy,K., Vidaud,D., Sobel,A. and Maucuer,A.
  TITLE     Quantitative RT-PCR reveals a ubiquitous but preferentially neural
            expression of the KIS gene in rat and human
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2909)
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-2003) Maucuer A., U440, Inserm, IFM, 17 rue du
            fer a moulin, 75005 Paris, FRANCE
FEATURES             Location/Qualifiers
     source          1. .2909
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:IMAGp998M213680"
                     /db_xref="taxon:9606"
                     /clone="I.M.A.G.E. Consortium CloneID 1455548"
                     /tissue_type="2 pooled tumors (clear cell type)"
     gene            1. .2909
                     /gene="KIS"
     CDS             146. .1405
                     /gene="KIS"
                     /codon_start=1
                     /product="KIS protein"
                     /protein_id="CAD60192.1"
                     /db_xref="GI:27657361"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN
```

Alignment Scores:

Pred. No.:	6.68e-188	Length:	2909
Score:	2195.00	Matches:	417
Percent Similarity:	99.76%	Conservative:	1
Best Local Similarity:	99.52%	Mismatches:	1
Query Match:	99.50%	Indels:	0
DB:	9	Gaps:	0

US-10-798-532-4 (1-419) x HSA536197 (1-2909)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 205

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 265

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||
Db        266 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 325

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 385

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 445

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 505

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 565

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 625

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 685

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 745
```

```
Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 805

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 865

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 925

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 985

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1045

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1105

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1106 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1165

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1225

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1285

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1345

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1346 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1402
```

```
RESULT 5
CQ412057
LOCUS       CQ412057                4065 bp    DNA     linear   PAT 23-JAN-2004
DEFINITION  Sequence 19128 from Patent WO0170979.
ACCESSION   CQ412057
VERSION     CQ412057.1  GI:41319838
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lee,J. and Lillie,J.
  TITLE     Genes, compositions, kits, and method for identification,
            assessment, prevention, and therapy of ovarian cancer
  JOURNAL   Patent: WO 0170979-A 19128 27-SEP-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .4065
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.07e-187       Length:         4065
Score:                  2195.00         Matches:        417
Percent Similarity:     99.76%          Conservative:   1
Best Local Similarity:  99.52%          Mismatches:     1
```

```
Query Match:            99.50%          Indels:         0
DB:                     6               Gaps:           0

US-10-798-532-4 (1-419) x CQ412057 (1-4065)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287
```

```
Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1524

RESULT 6
AX188296
LOCUS       AX188296               4065 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 3991 from Patent WO0142467.
ACCESSION   AX188296
VERSION     AX188296.1  GI:15139769
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
  TITLE     Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
  JOURNAL   Patent: WO 0142467-A 3991 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .4065
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                1.07e-187      Length:          4065
Score:                    2195.00        Matches:         417
Percent Similarity:       99.76%         Conservative:    1
Best Local Similarity:    99.52%         Mismatches:      1
Query Match:              99.50%         Indels:          0
DB:                       6              Gaps:            0

US-10-798-532-4 (1-419) x AX188296 (1-4065)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 327

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 387

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        388 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 447

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 507

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 567

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 627

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 687

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 747

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 807

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 867

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 927

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 987

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 1047

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1107

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1167

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1227

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1287

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1347

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1407

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1467

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1524

RESULT 7
AX166543
LOCUS       AX166543               1260 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 34 from Patent WO0138503.
ACCESSION   AX166543
VERSION     AX166543.1  GI:14546888
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
```

```
  AUTHORS   Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
            Flanagan,P. and Clary,D.S.
  TITLE     Novel human protein kinases and protein kinase-like enzymes
  JOURNAL   Patent: WO 0138503-A 34 31-MAY-2001;
            Sugen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1260
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              8.88e-188      Length:        1260
Score:                  2188.00        Matches:       416
Percent Similarity:     99.52%         Conservative:  1
Best Local Similarity:  99.28%         Mismatches:    2
Query Match:            99.18%         Indels:        0
DB:                     6              Gaps:          0

US-10-798-532-4 (1-419) x AX166543 (1-1260)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 60

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 120

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        121 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 180

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 240

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 300

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 360

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 420

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 480

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 540

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 660

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 720

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 780

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 840

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 900

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 960

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db        961 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1020

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1080

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1140

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1200

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1257

RESULT 8
AX680136
LOCUS       AX680136                1824 bp    DNA     linear   PAT 29-MAR-2003
DEFINITION  Sequence 21 from Patent WO0181555.
ACCESSION   AX680136
VERSION     AX680136.1  GI:29369916
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Gururajan,R., Patterson,C., Recipon,S.A., Tribouley,C.M.,
            Hafalia,A., Khan,F.Y., Yue,H., Au-Young,J., Bandman,O.,
            Baughn,M.R., Borowsky,M.L., Burford,N., Burrill,J.D., Elliot,V.S.,
            Gandhi,A.R., Kearney,L., Lal,P., Lu,D.A., Lu,Y., Tang,Y.T.,
            Zingler,K.A., Griffin,J.A., Hillman,J.J., Marcus,G.A., Nguyen,D.B.,
            Policky,J., Ramkumar,J., Thornton,M., Walia,N.K. and Walsh,R.T.
  TITLE     Human kinases
  JOURNAL   Patent: WO 0181555-A 21 01-NOV-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1824
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 3107952CB1"
ORIGIN

Alignment Scores:
Pred. No.:              1.49e-187      Length:        1824
Score:                  2188.00        Matches:       416
Percent Similarity:     99.52%         Conservative:  1
Best Local Similarity:  99.28%         Mismatches:    2
Query Match:            99.18%         Indels:        0
DB:                     6              Gaps:          0

US-10-798-532-4 (1-419) x AX680136 (1-1824)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
```

```
Db        169 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 228

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40

Db        229 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 288

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60

Db        289 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 348

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80

Db        349 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 408

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100

Db        409 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 468

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120

Db        469 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 528

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140

Db        529 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 588

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160

Db        589 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 648

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180

Db        649 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 708

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200

Db        709 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 768

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220

Db        769 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 828

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240

Db        829 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 888

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260

Db        889 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAC 948

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280

Db        949 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 1008

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300

Db       1009 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1068

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320

Db       1069 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1128

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340

Db       1129 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1188

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360

Db       1189 GAAGATGTTGTAGAAGATGTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTA 1248

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

Db       1249 CTTGTTCCAAAGGGAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGT 1308

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400

Db       1309 GATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTG 1368

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419

Db       1369 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1425
```

RESULT 9
RNKISRNA

LOCUS	RNKISRNA 1633 bp mRNA linear ROD 08-JUL-2002
DEFINITION	R.norvegicus mRNA for KIS protein.
ACCESSION	X98374
VERSION	X98374.1 GI:1403531
KEYWORDS	kis gene.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1
AUTHORS	Maucuer,A., Ozon,S., Manceau,V., Gavet,O., Lawler,S., Curmi,P. and Sobel,A.
TITLE	KIS is a protein kinase with an RNA recognition motif
JOURNAL	J. Biol. Chem. 272 (37), 23151-23156 (1997)
MEDLINE	97435279
PUBMED	9287318
REFERENCE	2 (bases 1 to 1633)
AUTHORS	Maucuer,A.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUN-1996) A. Maucuer, INSERM U440, 17 rue du Fer Moulin, 75005 Paris, FRANCE

FEATURES Location/Qualifiers

```
     source          1. .1633
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /tissue_type="brain"
                     /clone_lib="7"
                     /dev_stage="neonate"
     gene            1. .1633
                     /gene="KIS"
     CDS             39. .1298
                     /gene="KIS"
                     /codon_start=1
                     /protein_id="CAA67021.1"
                     /db_xref="GI:1403532"
                     /db_xref="GOA:Q63285"
                     /db_xref="Swiss-Prot:Q63285"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
```

ORIGIN

Alignment Scores:			
Pred. No.:	3.61e-187	Length:	1633
Score:	2183.00	Matches:	414
Percent Similarity:	99.52%	Conservative:	3
Best Local Similarity:	98.81%	Mismatches:	2
Query Match:	98.96%	Indels:	0
DB:	10	Gaps:	0

US-10-798-532-4 (1-419) x RNKISRNA (1-1633)

```
Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
```

```
Db        39 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 98

Qy        21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40

Db        99 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 158

Qy        41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60

Db       159 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 218

Qy        61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80

Db       219 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 278

Qy        81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100

Db       279 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 338

Qy       101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120

Db       339 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 398

Qy       121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140

Db       399 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 458

Qy       141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160

Db       459 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 518

Qy       161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180

Db       519 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 578

Qy       181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200

Db       579 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 638

Qy       201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220

Db       639 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 698

Qy       221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240

Db       699 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 758

Qy       241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260

Db       759 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 818

Qy       261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280

Db       819 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 878

Qy       281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300

Db       879 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 938

Qy       301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320

Db       939 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 998

Qy       321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340

Db       999 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1058

Qy       341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360

Db      1059 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1118

Qy       361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

Db      1119 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1178

Qy       381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400

Db      1179 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1238

Qy       401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419

Db      1239 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 1295
```

```
RESULT 10
RNU70372
LOCUS       RNU70372                3244 bp    mRNA    linear   ROD 08-SEP-1999
DEFINITION  Rattus norvegicus PAM COOH-terminal interactor protein 2 mRNA,
            complete cds.
ACCESSION   U70372
VERSION     U70372.2  GI:5821767
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 3244)
  AUTHORS   Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
            and Eipper,B.A.
  TITLE     Novel proteins that interact with the COOH-terminal cytosolic
            routing determinants of an integral membrane peptide-processing
            enzyme
  JOURNAL   J. Biol. Chem. 271 (45), 28636-28640 (1996)
  MEDLINE   97067094
   PUBMED   8910496
REFERENCE   2  (bases 1 to 3244)
  AUTHORS   Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
            Eipper,B.A. and Mains,R.E.
  TITLE     The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
            Determinants of the Peptide Processing Enzyme Peptidylglycine
            alpha-Amidating Monooxygenase
  JOURNAL   J. Biol. Chem. 274 (1999) In press
REFERENCE   3  (bases 1 to 3244)
  AUTHORS   Alam,R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
            and Eipper,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
            School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
            21205, USA
REFERENCE   4  (bases 1 to 3244)
  AUTHORS   Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
            Eipper,B.A. and Mains,R.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
            School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
            21205, USA
  REMARK    Sequence update by submitter
COMMENT     On Sep 3, 1999 this sequence version replaced gi:1698778.
FEATURES             Location/Qualifiers
     source          1. .3244
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
     CDS             162. .1421
                     /function="serine protein kinase"
                     /note="P-CIP2"
                     /codon_start=1
                     /product="PAM COOH-terminal interactor protein 2"
                     /protein_id="AAC53031.2"
                     /db_xref="GI:5821768"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLVYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
```

```
                    MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

Alignment Scores:
Pred. No.:              9.42e-187       Length:         3244
Score:                  2183.00         Matches:        414
Percent Similarity:     99.52%          Conservative:   3
Best Local Similarity:  98.81%          Mismatches:     2
Query Match:            98.96%          Indels:         0
DB:                     10              Gaps:           0

US-10-798-532-4 (1-419) x RNU70372 (1-3244)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 ATGGCCGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 221

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 281

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||   ||||||||||||||||||||||:::||||||||||||||||||||
Db        282 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 341

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 401

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 461

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 521

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              |||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATCCAGCACTGTGCCAGA 581

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GATGTCCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAG 641

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 701

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 761

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 821

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 881

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 941

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 1001

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 1061

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 1121

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||:::||||||
Db       1122 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 1181

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1182 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1241

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1242 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 1301

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1302 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1361

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1362 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 1418

RESULT 11
BC058732
LOCUS       BC058732                2041 bp    mRNA    linear   ROD 30-JUN-2004
DEFINITION  Mus musculus kinase interacting with leukemia-associated gene
            (stathmin), mRNA (cDNA clone MGC:64770 IMAGE:6414877), complete
            cds.
ACCESSION   BC058732
VERSION     BC058732.1  GI:37194892
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2041)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2041)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
           DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
           Thomas L. Casavant.
           Web site: http://genome.uiowa.edu
           Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
           Bonaldo,M.P., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
           Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
           Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,
           Casavant,T., Soares,M.B.

           Clone distribution: MGC clone distribution information can be found
           through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
           Series:  Plate:  Row:  Column: 0
           This clone was selected for full length sequencing because it
           passed the following selection criteria: matched mRNA gi: 6754447.
FEATURES             Location/Qualifiers
     source          1. .2041
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="MGC:64770 IMAGE:6414877"
                     /tissue_type="Brain, pool of enriched mouse brain tissues
                     from 13.5,14.5,16.5 and 17.5 dpc"
                     /clone_lib="NIH_BMAP_FR0"
                     /lab_host="DH10B"
                     /note="Vector: pYX-ASC"
     gene            1. .2041
                     /gene="Kist"
                     /note="synonym: KIS"
                     /db_xref="LocusID:16589"
                     /db_xref="MGI:1341908"
     CDS             154. .1413
                     /gene="Kist"
                     /codon_start=1
                     /product="kinase interacting with leukemia-associated gene
                     (stathmin)"
                     /protein_id="AAH58732.1"
                     /db_xref="GI:37194893"
                     /db_xref="LocusID:16589"
                     /db_xref="MGI:1341908"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYLYQTLL"
ORIGIN

Alignment Scores:
Pred. No.:              6.07e-187      Length:          2041
Score:                  2182.00        Matches:         414
Percent Similarity:     99.28%         Conservative:    2
Best Local Similarity:  98.81%         Mismatches:      3
Query Match:            98.91%         Indels:          0
DB:                     10             Gaps:            0

US-10-798-532-4 (1-419) x BC058732 (1-2041)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 213

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 273

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||   |||||||||||||||||||||||:::||||||||||||||||||
Db        274 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 333

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 393

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 453

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 513

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 573

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 633

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        634 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 693

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        694 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 753

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        754 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 813

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        814 TCAGCTGTTGATCTGTGGAGCCTCGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAG 873

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        874 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAT 933

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        934 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 993

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        994 ATCAAAAGCATGCTTCATGATGACCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGC 1053

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1054 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACT 1113

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||   |||||||:::||||||
Db       1114 CCAGTGCTCAGGCTCCTCAATGTGCTGGATGATGATTACCTTGAAAATGAAGATGAATAT 1173

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1174 GAAGATGTTGTAGAAGATGTGAAAGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1233

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1234 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1293

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1294 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1353

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1354 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 1410

RESULT 12
AC129141_1/c
```

```
WPCOMMENT
  Sequence split into 4 fragments  LOCUS AC129141 Accession AC129141
     Fragment Name      Begin     End
     AC129141_0             1     110000
     AC129141_1        100001     210000
     AC129141_2        200001     310000
     AC129141_3        300001     363206
Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH2

Alignment Scores:
Pred. No.:              8.44e-184       Length:         110000
Score:                  2174.00         Matches:        413
Percent Similarity:     99.28%          Conservative:   3
Best Local Similarity:  98.57%          Mismatches:     3
Query Match:            98.55%          Indels:         0
DB:                     2               Gaps:           0

US-10-798-532-4 (1-419) x AC129141_1 (1-110000)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              |||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db      70316 ATGGCCGGGTCCTGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 70257

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70256 CGGCTGTGGCAGGTACAGAGCCGCCTAGGCAGCGGCTCCTCGGCCTCGGTGTACCGGGTG 70197

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              |||||||||||   |||||||||||||||||||:::||||||||||||||||||||||||
Db      70196 CGCTGCTGCGGCACTCCAGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 70137

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70136 ACCACCGGGGCTGCCGCCTCGGCTGCGGAGTATGGTTTCCGAAAAGAGAGGGCAGCGCTG 70077

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70076 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATACGGAGTCTTTACCATACACTTC 70017

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70016 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGCGTTTCGGAA 69957

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              |||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69956 TTGCTTGTATATTCCAGTCACCAGGGTTGCTCCATGTGGATGATACAGCACTGTGCCAGA 69897

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69896 GATGTTCTGGAGGCCCTTGCTTTTCTTCACCATGAGGGCTATGTCCATGCAGACCTCAAA 69837

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69836 CCACGAAACATCCTGTGGAGTGCGGAGAACGAGTGCTTTAAGCTTATTGACTTTGGACTC 69777

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69776 AGCTTCAAAGAAGGCAATCAGGACGTGAAGTATATTCAGACAGACGGGTATCGAGCTCCT 69717

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69716 GAAGCAGAACTGCAGAATTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAGTGTACC 69657

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69656 TCAGCTGTTGATCTCTGGAGCCTTGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 69597

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69596 CTGAAACATACAGTCAGATCTCAGGAGTGGAAGGCAAACAGTTCTGCTATTATTGATCAT 69537

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69536 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 69477

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69476 ATCAAAAGCATGCTTCATGACGACCCAAGCAGAAGGATCCCTGCTGAGATGGCCTTGTGC 69417

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69416 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCAACT 69357

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              |||||||||||||||||||||||||||||||||||||||||||   ||||||:::||||||
Db      69356 CCAGTGCTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAAAATGAAGATGAATAT 69297

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69296 GAAGATGTTGTAGAAGATGTAAAAGAAGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 69237

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69236 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTTGTTGAGTACGCAAACGCTGGA 69177

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69176 GATTCCAAAGCTGCTCAGAAATTGCTGACTGGGAGGATGTTTGATGGGAAGTTTGTTGTG 69117

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69116 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTTTATCAAACCTTGCTT 69060

RESULT 13
MMPKKIS
LOCUS       MMPKKIS                 1254 bp    mRNA    linear   ROD 08-JUL-2002
DEFINITION  Mus musculus partial mRNA for KIS protein kinase (Kis gene).
ACCESSION   Y10725
VERSION     Y10725.2  GI:21726712
KEYWORDS    Kis gene; protein kinase.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1254)
  AUTHORS   Maucuer,A., Ozon,S., Manceau,V., Gavet,O., Lawler,S., Curmi,P. and
            Sobel,A.
  TITLE     KIS is a protein kinase with an RNA recognition motif
  JOURNAL   J. Biol. Chem. 272 (37), 23151-23156 (1997)
  MEDLINE   97435279
   PUBMED   9287318
REFERENCE   2
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-1997) Maucuer A., INSERM U440, U440, 17 Rue du
            Fer Moulin, Paris, 75005, FRANCE
  REMARK    revised by [4]
REFERENCE   3  (bases 1 to 1254)
  AUTHORS   Maucuer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) A. Maucuer, INSERM U440, U440, 17 Rue du
            Fer Moulin, Paris, 75005, FRANCE
COMMENT     On Jul 10, 2002 this sequence version replaced gi:1806129.
            Overlaps with related sequences X82320, X10725.
FEATURES             Location/Qualifiers
     source          1. .1254
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="K22"
                     /tissue_type="brain"
                     /clone_lib="lambda gt10"
                     /dev_stage="neonate"
     gene            1. .1254
                     /gene="Kis"
     CDS             11. .>1254
```

```
                     /gene="Kis"
                     /codon_start=1
                     /product="KIS protein kinase"
                     /protein_id="CAA71714.2"
                     /db_xref="GI:21726713"
                     /db_xref="GOA:P97343"
                     /db_xref="Swiss-Prot:P97343"
                     /translation="MAGSGCAWGAEPPRFLEAFGRLWQVQSRLGSGSSASVYRVRCCG
                     TPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSAIIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPGRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEDEYEDVVEDVKEECQKYGPVVSLLVPKENPGRGQVFVEYANAGDSKAAQKLLTGR
                     MFDGKFVVATFYPLSAYKRGYL"
ORIGIN

Alignment Scores:
Pred. No.:                5.55e-185      Length:          1254
Score:                    2157.00        Matches:         409
Percent Similarity:       99.28%         Conservative:    2
Best Local Similarity:    98.79%         Mismatches:      3
Query Match:              97.78%         Indels:          0
DB:                       10             Gaps:            0

US-10-798-532-4 (1-419) x MMPKKIS (1-1254)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ATGGCGGGGTCCGGCTGCGCGTGGGGCGCCGAGCCGCCGCGCTTCCTGGAGGCCTTCGGG 70

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 CGGCTGTGGCAGGTCCAGAGCCGCCTGGGCAGCGGCTCGTCGGCCTCGGTGTACCGGGTG 130

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              |||||||||||   |||||||||||||||||||||||:::|||||||||||||||||||
Db        131 CGGTGCTGCGGTACCCCGGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCCTGCCTCCGGGA 190

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 ACCACCGGGGCTGCGGCCTCGGCCGCGGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 250

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 GAGCAGTTGCAGGGTCACAGGAACATCGTGACTTTATATGGAGTCTTTACCATACACTTC 310

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 TCTCCCAATGTGCCATCACGCTGTCTGCTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 370

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 TTGCTTTTATATTCCAGTCATCAGGGCTGCTCCATGTGGATGATACAGCACTGTGCCCGA 430

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 GATGTTCTGGAGGCCCTCGCTTTTCTTCACCATGAGGGCTACGTCCATGCAGACCTCAAA 490

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 CCACGAAACATCCTGTGGAGTGCGGAGAATGAGTGTTTTAAGCTTATTGACTTTGGACTC 550

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 AGCTTCAAAGAAGGCAATCAGGACGTAAAGTATATTCAGACAGACGGGTATCGCGCTCCT 610

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 GAAGCAGAGCTGCAGAACTGCTTGGCCCAGGCCGGCCTGCAGAGTGATACAGAATGTACC 670

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        671 TCAGCTGTTGATCTGTGGAGCCTCGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAG 730

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        731 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCAT 790

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        791 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTCAGAGACCTT 850

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        851 ATCAAAAGCATGCTTCATGATGACCCGGGCAGAAGAATCCCTGCTGAGATGGCATTGTGC 910

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        911 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTGATGCTTCCGACT 970

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        971 CCAGTGCTCAGGCTCCTCAATGTGCTGGATGATGATTACCTTGAAAATGAAGATGAATAT 1030

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1031 GAAGATGTTGTAGAAGATGTGAAAGAGGAGTGTCAGAAATATGGACCAGTGGTTTCTCTG 1090

Qy        361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 CTTGTTCCAAAGGAAAATCCTGGCAGAGGACAAGTCTTCGTTGAGTACGCGAACGCTGGT 1150

Qy        381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1151 GATTCCAAAGCTGCTCAGAAGTTGCTGACTGGGAGGATGTTTGACGGGAAGTTTGTTGTG 1210

Qy        401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeu 414
              |||||||||||||||||||||||||||||||||||||||||||
Db       1211 GCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTT 1252

RESULT 14
BC026046
LOCUS       BC026046               1367 bp    mRNA    linear   PRI 30-JUN-2004
DEFINITION  Homo sapiens kinase interacting with leukemia-associated gene
            (stathmin), mRNA (cDNA clone MGC:39391 IMAGE:4649261), complete
            cds.
ACCESSION   BC026046
VERSION     BC026046.1  GI:19684094
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1367)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
```

```
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1367)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
            Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
            Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
            Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
            Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,
            Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skaleka,
            Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
            Schein, Asim Siddiqui, Rob Holt, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 42 Row: f Column: 6
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 21389438.
FEATURES             Location/Qualifiers
     source          1. .1367
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:39391 IMAGE:4649261"
                     /tissue_type="Eye, retinoblastoma"
                     /clone_lib="NIH_MGC_16"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1. .1367
                     /gene="KIS"
                     /db_xref="LocusID:127933"
     CDS             159. .1193
                     /gene="KIS"
                     /codon_start=1
                     /product="kinase interacting stathmin"
                     /protein_id="AAH26046.1"
                     /db_xref="GI:19684095"
                     /db_xref="LocusID:127933"
                     /translation="MAGSGCAWGAEPPRFLEAPGRLWQVQSRLGSGSSASVYRVRCCG
                     NPGSPPGALKQFLPPGTTGAAASAAEYGFRKERAALEQLQGHRNIVTLYGVFTIHFSP
                     NVPSRCLLLELLDVSVSELLLYSSHQGCSMWMIQHCARDVLEALAFLHHEGYVHADLK
                     PRNILWSAENECFKLIDFGLSFKEGNQDVKYIQTDGYRAPEAELQNCLAQAGLQSDTE
                     CTSAVDLWSLGIILLEMFSGMKLKHTVRSQEWKANSSATIDHIFASKAVVNAAIPAYH
                     LRDLIKSMLHDDPSRRIPAEMALCSPFFSIPFAPHIEDLVMLPTPVLRLLNVLDDDYL
                     ENEEEYEGLC"
ORIGIN

Alignment Scores:
Pred. No.:              7.45e-170       Length:         1367
Score:                  1990.00         Matches:        385
Percent Similarity:     92.14%          Conservative:   2
Best Local Similarity:  91.67%          Mismatches:     2
Query Match:            90.21%          Indels:         31
DB:                     9               Gaps:           1

US-10-798-532-4 (1-419) x BC026046 (1-1367)

Qy          1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 ATGGCGGGATCCGGCTGCGCCTGGGGCGCGGAGCCGCCGCGTTTTCTGGAGGCCTTCGGG 218

Qy         21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTT 278

Qy         41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
              ||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db        279 CGCTGCTGCGGCAACCCTGGCTCGCCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGA 338

Qy         61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 ACCACCGGGGCTGCGGCCTCTGCCGCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTG 398

Qy         81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 GAACAGTTGCAGGGTCACAGAAACATCGTGACTTTGTATGGAGTGTTTACAATCCACTTT 458

Qy        101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 TCTCCAAATGTGCCATCACGCTGTCTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAA 518

Qy        121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 TTGCTCTTATATTCCAGTCACCAGGGTTGTTCCATGTGGATGATACAGCATTGTGCCCGA 578

Qy        141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGCTATGTCCATGCGGACCTCAAA 638

Qy        161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 CCACGTAACATATTGTGGAGTGCAGAGAATGAATGTTTTAAACTCATTGACTTTGGACTT 698

Qy        181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 AGCTTCAAAGAAGGCAATCAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCA 758

Qy        201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 GAAGCAGAATTGCAAAATTGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACC 818

Qy        221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        819 TCAGCTGTTGATCTGTGGAGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 878

Qy        241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAAACAGTTCTGCTACTATTGATCAC 938

Qy        261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ATATTTGCCAGTAAAGCAGTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTT 998

Qy        281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 ATCAAAAGCATGCTTCATGATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGC 1058

Qy        301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AGCCCATTCTTTAGCATTCCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1118

Qy        321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyr 340
              ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db       1119 CCAGTGCTAAGACTGCTGAATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATAT 1178

Qy        341 GluAspValValGluAspValLysGluGluCysGlnLysTyrGlyProValValSerLeu 360

Db       1178 ------------------------------------------------------------ 1178

Qy        361 LeuValProLysGluAsnProGlyArgGlyGln-ValPheValGluTyrAlaAsnAlaGl 380
                                                  ::: ||||||||||||||||||||||||||
Db       1179 ----------------------------GAAGGTCTTTGTTGAGTATGCAAATGCTGG 1208
```

```
Qy        380 yAspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVa 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1209 TGATTCCAAAGCTGCGCAGAAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGT 1268

Qy        400 lAlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 GGCTACATTCTACCCGCTGAGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1326

RESULT 15
HSM805327
LOCUS       HSM805327               2940 bp    mRNA    linear   PRI 12-JUL-2002
DEFINITION  Homo sapiens mRNA; cDNA DKFZp667I0511 (from clone DKFZp667I0511).
ACCESSION   AL834136
VERSION     AL834136.1  GI:21739600
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2940)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp667I0511) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .2940
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:DKFZp667I0511"
                     /db_xref="taxon:9606"
                     /map="1"
                     /clone="DKFZp667I0511"
                     /tissue_type="lymph node"
                     /clone_lib="667 (synonym: hlno2). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     polyA_site      2923
ORIGIN

Alignment Scores:
Pred. No.:              2.09e-160       Length:         2940
Score:                  1890.50         Matches:        388
Percent Similarity:     63.46%          Conservative:   1
Best Local Similarity:  63.30%          Mismatches:     2
Query Match:            85.70%          Indels:         223
DB:                     9               Gaps:           1

US-10-798-532-4 (1-419) x HSM805327 (1-2940)

Qy         29 LeuGlySerGlySerSerAlaSerValTyrArgValArgCysCysGlyAsnProGlySer 48
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGGGTAGCGGCTCCTCCGCCTCGGTGTATCGGGTTCGCTGCTGCGGCAACCCTGGCTCG 60

Qy         49 ProProGlyAlaLeuArgGlnPheLeuProProGlyThrThrGlyAlaAlaAlaSerAla 68
              ||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCCCCGGCGCCCTCAAGCAGTTCTTGCCGCCAGGAACCACCGGGGCTGCGGCCTCTGCC 120

Qy         69 AlaGluTyrGlyPheArgLysGluArgAlaAlaLeuGluGlnLeuGlnGlyHisArgAsn 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCGAGTATGGTTTCCGCAAAGAGAGGGCGGCGCTGGAACAGTTGCAGGGTCACAGAAAC 180

Qy         89 IleValThrLeuTyrGlyValPheThrIleHisPheSerProAsnValProSerArgCys 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATCGTGACTTTGTATGGAGTGTTTACAATCCACTTTTCTCCAAATGTGCCATCACGCTGT 240

Qy        109 LeuLeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuTyrSerSerHisGln 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGTTGCTTGAACTCCTGGATGTCAGTGTTTCGGAATTGCTCTTATATTCCAGTCACCAG 300

Qy        129 GlyCysSerMetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPhe 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTGTTCCATGTGGATGATACAGCATTGTGCCCGAGATGTTTTGGAGGCCCTTGCTTTT 360

Qy        149 LeuHisHisGluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAla 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTCATCATGAGGGCTATGTCCATGCGGACCTCAAACCACGTAACATATTGTGGAGTGCA 420

Qy        169 GluAsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGln--- 187
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGAATGAATGTTTTAAACTCATTGACTTTGGACTTAGCTTCAAAGAAGGCAATCA-GGT 479

Qy        187 ------------------------------------------------------------ 187

Db        480 AAGAAATAACCTTTTCTTTTCTCTCGCAGTTTAAAATGTAGTGGTTGAAGCCAAATATTT 539

Qy        187 ------------------------------------------------------------ 187

Db        540 TTAAATGACAAGGGTAGTTTAGTAGGGGCTTACAAGTCATTTTAACCTAAATTACTGAAA 599

Qy        187 ------------------------------------------------------------ 187

Db        600 TGCCAGGGGATGCAGTAAATCAAAAGAATTTCAGTGCCACCGTGTAAAAAAAGGACATAA 659

Qy        187 ------------------------------------------------------------ 187

Db        660 TGTTTGGGTCAGGTAGGAGGAGCCAGTGAAGCATGAGAGCTCTGGAATTATTAAAAGTAC 719

Qy        187 ------------------------------------------------------------ 187

Db        720 TTAACCTACATAGAATGGTTTTTTAAATTTTATGCGTGTTTGGCTAATTGCTGAGTTACT 779

Qy        187 ------------------------------------------------------------ 187

Db        780 ATTCCTGGTCTCTTCAGTGTTGGGGCTGTATCTCGTGGCAGAATGCATATGATAGTTCTG 839

Qy        187 ------------------------------------------------------------ 187

Db        840 CCAAATCCTCTTTTCTCCATGGTTTATCATTTCTTTGCAGTGTCCTTATTATCAGCAGTG 899

Qy        187 ------------------------------------------------------------ 187

Db        900 CCTAATAGGTAAGAGACAATCCAAAGAAGACTTCTTTATGAGCGCTATATGGAAATGTAC 959

Qy        187 ------------------------------------------------------------ 187

Db        960 AGCTTGCCTCTGAGAACTGCCCAGTTAATGAACCAAAGAATTATGGTAATCAAGGACTAA 1019

Qy        187 ------------------------------------------------------------ 187

Db       1020 AAATTTTTAGCATTTAGCTTGGCAGTGGGTTTACTGCACTCTTAGGGAAGTAAATTCACT 1079

Qy        187 ------------------------------------------------------------ 187

Db       1080 GTAGAGATGGATAGAAAGGGAGCAGCTTTTTTATTGGTTGTAATATTTACTCATCTTTTT 1139

Qy        188 ---AspValLysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsn 206
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 TAGGATGTAAAGTATATTCAGACAGACGGGTATCGGGCTCCAGAAGCAGAATTGCAAAAT 1199

Qy        207 CysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrp 226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 TGCTTGGCCCAGGCTGGCCTGCAGAGTGATACAGAATGTACCTCAGCTGTTGATCTGTGG 1259

Qy        227 SerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArg 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 AGCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAGA 1319
```

```
Qy        247 SerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAla 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 TCTCAGGAATGGAAGGCAAACAGTTCTGCTATTATTGATCACATATTTGCCAGTAAAGCA 1379

Qy        267 ValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHis 286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 GTGGTGAATGCCGCAATTCCAGCCTATCACCTAAGAGACCTTATCAAAAGCATGCTTCAT 1439

Qy        287 AspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIle 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 GATGATCCAAGCAGAAGAATTCCTGCTGAAATGGCATTGTGCAGCCCATTCTTTAGCATT 1499

Qy        307 ProPheAlaProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeu 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 CCTTTTGCCCCTCATATTGAAGATCTGGTCATGCTTCCCACTCCAGTGCTAAGACTGCTG 1559

Qy        327 AsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAsp 346
              ||||||||||||||||||||||    ||||||||||||||||||||||||||||||||||
Db       1560 AATGTGCTGGATGATGATTATCTTGAGAATGAAGAGGAATATGAAGATGTTGTAGAAGAT 1619

Qy        347 ValLysGluGluCysGlnLysTyrGlyProValValSerLeuLeuValProLysGluAsn 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 GTAAAAGAGGAGTGTCAAAAATATGGACCAGTGGTATCTCTACTTGTTCCAAAGGAAAAT 1679

Qy        367 ProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGlyAspSerLysAlaAlaGln 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 CCTGGCAGAGGACAAGTCTTTGTTGAGTATGCAAATGCTGGTGATTCCAAAGCTGCGCAG 1739

Qy        387 LysLeuLeuThrGlyArgMetPheAspGlyLysPheValValAlaThrPheTyrProLeu 406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 AAATTACTGACTGGAAGGATGTTTGATGGGAAGTTTGTTGTGGCTACATTCTACCCGCTG 1799

Qy        407 SerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
              |||||||||||||||||||||||||||||||||||||||
Db       1800 AGTGCCTACAAGAGGGGATATCTGTATCAAACCTTGCTT 1838
```

Search completed: November 29, 2004, 18:11:29
Job time : 4618 secs